



THÈSE / UNIVERSITÉ DE RENNES 1
sous le sceau de l'Université Européenne de Bretagne

pour le grade de

DOCTEUR DE L'UNIVERSITÉ DE RENNES 1

Mention : TRAITEMENT DU SIGNAL ET TÉLÉCOMMUNICATIONS

École doctorale MATISSE

présentée par

Chunfeng YANG

préparée à l'unité de recherche LTSI - INSERM UMR 1099
Laboratoire Traitement du Signal et de l'Image
UFR Informatique Electronique (ISTIC)

**Contribution
à l'analyse de
connectivité
effective
en épilepsie**

**Thèse soutenue à Rennes
le 10 juillet 2012**

devant le jury composé de :

Valérie LOUIS DORR

Professeur des Universités / *rapporteur*

Jacques DUCHENE

Professeur des Universités / *rapporteur*

Jean-Marc EDELINE

Directeur de Recherche CNRS / *président du jury*

Limin LUO

Professeur des Universités / *examineur*

Huazhong SHU

Professeur des Universités / *examineur*

Jean-Jacques BELLANGER

Maître de Conférences / *examineur*

Régine LE BOUQUIN JEANNES

Professeur des Universités / *directeur de thèse*

Remerciements

Cette thèse s'inscrit dans le cadre du CRIBs (Centre de Recherche en Information Biomédicale sino-français), Laboratoire International Associé à l'INSERM, à l'Université de Rennes 1 et à l'Université du Sud-Est à Nankin.

Tout d'abord, je souhaite vivement remercier ma directrice de thèse, le Professeur Régine LE BOUQUIN JEANNES. Sa rigueur et sa méticulosité dans sa recherche, sa passion, son abnégation dans le travail, et son exceptionnelle vigueur m'ont extrêmement impressionné et resteront pour moi un modèle et un exemple à suivre dans ma future carrière. J'exprimerai également ma reconnaissance à Jean-Jacques BELLANGER pour les discussions riches et fructueuses que nous avons eues au cours de ces années et pour son aide dans le développement de méthodes ainsi que pour son investissement lors de la préparation de ma soutenance. Je remercie Régine et Jean-Jacques pour leur aide précieuse et pour le temps qu'ils ont passé sur ma thèse (de jour comme de nuit, week-ends inclus !). Je voudrais témoigner toute ma gratitude au Professeur Gérard FAUCON pour l'intérêt qu'il a porté à ce travail lors de son initiation et pour nos nombreuses discussions.

J'adresse aussi mes remerciements au Professeur Limin LUO qui m'a initialement accueilli comme doctorant au sein de son laboratoire à l'Université du Sud-Est à Nankin, le LIST (Laboratory of Image and Science Technology) et qui m'a donné la chance de préparer mon doctorat en France. Je tiens également à remercier sincèrement le Professeur Huazhong SHU qui m'a assuré son appui tout au long de ma thèse et pour ses suggestions. Je suis reconnaissant au Professeur Lotfi SENHADJI pour m'avoir accueilli dans son laboratoire, le LTSI (Laboratoire Traitement du Signal et de l'Image) à l'Université de Rennes 1 ainsi qu'à Fabrice WENDLING pour ses conseils avisés particulièrement sur le modèle physiologique.

J'ai été très honoré que les Professeurs Valérie LOUIS DORR et Jacques DUCHENE acceptent la tâche de rapporteur et que Jean-Marc EDELINE préside le jury lors de ma soutenance.

Je ne voudrais surtout pas oublier de remercier toutes les personnes qui m'ont aidé et avec lesquelles j'ai travaillé toutes ces années, que ce soit au sein du LTSI entre 2008 et 2012 ou au LIST depuis 2007.

D'autre part, j'ai beaucoup apprécié Régine et sa famille (Eric, Thomas, Juline and Marie), ainsi que Gérard et sa femme Marie-Jeanne, pour leur gentillesse et leur hospitalité qui m'ont

permis de développer ma recherche dans une ambiance et un climat chaleureux, loin de ma famille.

Naturellement, j'ai une pensée toute particulière pour mon épouse, Xiaojia WANG, et nos parents (Guochun YANG, Chanjuan DU, Zhiming WANG and Yuefen HUANG) ainsi que pour mon futur bébé. Je les remercie pour leur bienveillance et leur soutien sans faille dans les moments les plus difficiles. Ils m'ont toujours supporté et encouragé, et leur amour sans limite a permis l'aboutissement de ce travail.

Résumé

Les travaux présentés dans cette thèse s'inscrivent dans la problématique de l'identification de la connectivité effective, et, de ce fait, de graphes de propagation, pour détecter et quantifier les relations entre structures cérébrales impliquées lors de l'initiation et de la diffusion de crises d'épilepsie. Définie comme la possibilité de connexion offerte par une entité à des entités voisines et/ou distantes, la connectivité se décline au niveau cérébral suivant trois modes : la connectivité anatomique ou structurelle fait référence aux liens physiques qui existent entre différentes régions anatomiques du cerveau, la connectivité fonctionnelle permet de caractériser l'activité en réseau du cerveau et la connectivité effective complète les notions de connectivités structurelle et fonctionnelle en introduisant le concept d'influence causale exercée par un système neuronal sur un autre, soit directement soit indirectement. Dans cette optique, différentes approches ont été envisagées afin d'apporter des éléments de réponse à deux questions essentielles, à savoir l'identification de liaisons unilatérales et/ou bilatérales entre structures et la mise en évidence de liens directs et/ou indirects. Après avoir recensé les principales techniques de la littérature, nous nous sommes intéressés à l'indice de causalité de Granger ainsi qu'à des extensions fréquentielles et/ou conditionnées, avant d'explorer un indice de pente de phase, l'idée sous-jacente étant l'estimation de retard. Un nouvel indice construit sur la cohérence dirigée partielle a été proposé permettant non seulement de détecter des flux réciproques mais également de discriminer relations directes et indirectes entre signaux. Dans un troisième temps, nous nous sommes orientés vers des techniques plus complexes non paramétriques (à des horizons de prédiction près) relevant de la théorie de l'information et plus spécifiquement vers l'entropie de transfert. De par la subordination et la sensibilité de cette technique au choix de paramètres de calibration comme l'ordre des modèles de Markov, nous avons proposé une optimisation de l'estimation de l'ordre de ces modèles à partir du critère d'information bayésien avant de considérer une mesure d'entropie de transfert conditionnelle dans un souci d'analyse multivariée. Les différentes approches proposées ont été évaluées et comparées sur des signaux simulés suivant des processus vectoriels autorégressifs linéaires et non linéaires ainsi que sur des modèles physiologiques réalistes avant d'être appliquées sur des signaux réels enregistrés sur un modèle animal (cochon d'inde). En simulation, les résultats obtenus permettent d'établir des graphes de propagation cohérents et conformes aux modèles et, dans le cas réel, d'apprécier les variations de cette connectivité au cours du temps.

Table des matières

Introduction	1
Chapitre 1. Contexte clinique et problématique	5
1.1. Le cerveau humain.....	5
1.2. L'épilepsie	6
1.3. Vers quelles solutions ?	8
1.3.1. Les avancées scientifiques et médicales	8
1.3.2. Problématique de l'étude.....	9
Chapitre 2. Connectivité et modèles.....	11
2.1. Connectivité cérébrale	11
2.1.1. Généralités	12
2.1.1.1. Modes de connectivité	12
2.1.1.2. Le formalisme des graphes.....	13
2.1.2. Techniques d'observation et approches computationnelles	14
2.1.2.1. Connectivité structurelle.....	14
2.1.2.2. Connectivités fonctionnelle et effective.....	15
2.1.2.2.1. Modalités instrumentales.....	15
2.1.2.2.2. Traitement des données dans la connectivité fonctionnelle.....	15
2.1.2.2.2.1. Techniques linéaires	15
2.1.2.2.2.2. Techniques non linéaires	16
2.1.2.2.3. Traitement des données dans la connectivité effective	17
2.1.2.2.3.1. Techniques basées sur un modèle	17
2.1.2.2.3.2. Techniques pilotées par les données.....	18
2.1.3. Relations entre connectivités structurelle, fonctionnelle et effective	20
2.2. Modèles neurocomputationnels de l'épilepsie	21
2.2.1. Modèles microscopiques.....	21
2.2.2. Modèles macroscopiques.....	22
Chapitre 3. Protocole expérimental	25
3.1. Bases de données	27
3.1.1. Modèles linéaires autorégressifs.....	27
3.1.2. Modèles non linéaires autorégressifs.....	27
3.1.3. Modèles physiologiques	28
3.1.4. Signaux réels.....	29
3.2. Méthodologie d'évaluation	29
3.2.1. Détermination de graphes	29
3.2.2. Analyse des distributions sous H_0	30

Chapitre 4. Méthodes.....	33
4.1. Indice de causalité de Granger	34
4.2. Indices basés sur la pente de phase.....	34
4.3. Entropie de transfert et entropie de transfert conditionnée	36
4.4. Discussion	37
Chapitre 5. Résultats expérimentaux	39
5.1. Analyse des résultats sur modèles AR linéaires	40
5.2. Analyse des résultats sur modèles AR non linéaires	41
5.3. Analyse des résultats sur modèles physiologiques.....	45
5.4. Analyse des résultats sur signaux réels.....	45
5.5. Discussion	46
Conclusion	49
Bibliographie	53

Introduction

Les patients souffrant d'épilepsie réfractaire à tout traitement médicamenteux bénéficient d'une évaluation pré-chirurgicale visant à délimiter la zone épileptogène, ou plus généralement un réseau épileptogène, responsable de la génération et/ou de la propagation des crises. Celle-ci peut s'envisager via l'enregistrement de signaux intracérébraux recueillis sur des électrodes de profondeur. Si cette procédure est invasive, elle permet, en revanche, l'exploration plus fine d'une partie du territoire cérébral. Notre étude vise à contribuer méthodologiquement à l'exploration de ces données, soit, plus précisément, à l'extraction des caractéristiques du réseau épileptogène, par application de méthodes statistiques d'analyse de connectivité.

Pouvant être définie comme la possibilité de connexion offerte par une entité à des entités voisines et/ou distantes, la connectivité se décline au niveau cérébral suivant trois modes [Sporns 2007] :

- la connectivité anatomique : elle fait référence aux liens physiques ou structurels qui existent entre différentes régions anatomiques du cerveau. Le terme "structure" est pris au sens large et peut représenter des éléments tant microscopiques, *i.e.* à l'échelle du neurone (concept de connexion synaptique entre des neurones), que mésoscopiques, *i.e.* à l'échelle de populations neuronales (concept d'ensembles synaptiques ou de "paquets" de connexions entre des réseaux de colonnes corticales), voire macroscopiques, *i.e.* à l'échelle des régions cérébrales (concept d'inter-connexion suivant différents chemins ou voies entre des aires distinctes du cerveau). Par essence, la connectivité anatomique sous-tend la notion de complexité architecturale au niveau cérébral ;
- la connectivité fonctionnelle : se cache sous cette dénomination une caractérisation de l'activité en réseau du cerveau. Ce concept, qui se ramène à des idées statistiques simples, porte sur les variations de "collaboration" (de co-activation) et donc de dépendance (incluant évidemment la possibilité d'indépendance) entre des unités neuronales distribuées (ou des groupes neuronaux) et souvent spatialement distantes. L'étude de réseaux fonctionnels passe ainsi non seulement par l'étude d'interactions existant entre ces différentes entités mais aussi par l'évolution de ces interactions au cours du temps. La connectivité fonctionnelle diffère de la connectivité anatomique dans le sens où une connexion anatomique peut ou non, de manière non stationnaire, supporter un lien fonctionnel. De même, qui dit connectivité fonctionnelle entre

deux aires ne dit pas nécessairement existence d'une connectivité anatomique directe entre ces structures puisque leur interdépendance peut être le résultat de la médiation par une troisième structure, voire par un ensemble d'autres structures. La dépendance statistique peut être estimée par des mesures de corrélation, de covariance, de cohérence spectrale, de synchronisation de phase, d'information mutuelle. Généralement, cette connectivité ne se réfère explicitement à aucune notion de directivité ou de causalité. Seules des intensités de couplage sont prises en considération ;

- la connectivité effective : elle complète les notions de connectivités structurelle et fonctionnelle en introduisant le concept d'influence causale exercée par un système neuronal sur un autre, soit directement soit indirectement. Au-delà de la connectivité fonctionnelle qui traduit le fonctionnement coordonné (non indépendant) de deux systèmes, la connectivité effective introduit la notion de cause à effet, dans le sens où elle cherche à exprimer la dépendance fonctionnelle d'un système sous l'influence d'un autre. Cette notion renvoie à celle de sens de circulation ou de flux d'information si l'on considère le cerveau comme un organe de traitement réparti de l'information. En principe, les effets causaux peuvent être inférés au travers de perturbations systématiques du système ou, lorsque ces perturbations ne sont pas applicables, à travers une analyse de prédictibilité des séries temporelles observées (dans la mesure où, suivant le principe de causalité en physique, les causes doivent précéder leurs conséquences et permettre de les prédire, au moins partiellement). Pour appréhender cette connectivité, certaines techniques nécessitent la spécification d'un modèle de génération des observations pouvant se résumer par un ensemble de paramètres (structurels ou influant sur les dynamiques temporelles locales), d'autres techniques étant non paramétriques.

Formellement, un modèle de connectivité cérébrale peut être représenté sous forme d'un graphe de propagation. Ce graphe peut être pondéré, et orienté dans le cas de la connectivité effective. Les pondérations représentent l'importance des connexions entre les différents éléments impliqués. Les orientations recherchées en connectivité effective sont elles-mêmes représentées par des flèches unidirectionnelles ou bidirectionnelles en cas d'influences réciproques (on peut imaginer deux populations neuronales pouvant se stimuler en miroir via deux faisceaux nerveux distincts). Ainsi, la connectivité fonctionnelle forme une matrice, dans laquelle chaque élément code une dépendance statistique entre deux éléments du système, et qui est symétrique dans la mesure où l'indicateur de dépendance est lui-même symétrique (comme c'est le cas du coefficient de corrélation linéaire). Un seuillage peut être appliqué à une telle matrice pour conduire à des graphes binaires non orientés. Quant à la connectivité effective, elle conduit à une matrice le plus souvent non symétrique, et appliquer un seuil à cette matrice induit des graphes binaires orientés.

Cette thèse s'inscrit dans cette problématique de détermination de connectivité effective, et, de ce fait, de graphes de propagation, pour détecter et quantifier les relations entre structures cérébrales impliquées lors de l'initiation et de la diffusion de crises d'épilepsie. On y propose une évaluation de méthodes linéaires et non linéaires, dont certaines sont originales. Cette évaluation s'appuie, pour caractériser les algorithmes proposés, sur des simulations physiologiquement réalistes. Cette

évaluation est complétée, en amont, par des simulations sur modèles autorégressifs et, en aval, par quelques signaux réels, prélevés sur un modèle animal d'épilepsie.

Le travail présenté se situe en amont d'une analyse plus systématique sur une base suffisamment riche de signaux réels puisque les modèles physiologiques réalistes utilisés sont configurés pour simuler l'activité de populations neuronales impliquées lors du décours d'une crise. Il constitue une première étape dans l'identification et la localisation de modifications de connectivité à des fins de diagnostic en termes d'organisation du réseau épileptique.

Ce document est organisé de la façon suivante : le premier chapitre permet de se familiariser avec le contexte médical et rappelle quelques notions d'anatomie cérébrale avant de présenter la problématique clinique. Sans être sans doute complètement exhaustif, le chapitre 2 tente de synthétiser les principaux travaux issus de la littérature portant sur la connectivité effective ainsi que le rôle joué par les modèles neurocomputationnels. Dans le chapitre 3, nous décrivons le protocole expérimental de cette étude et commençons par décrire les modèles que nous avons retenus pour simuler des signaux de crise utilisés pour tester les méthodes et algorithmes que nous avons choisis d'investir, et développés plus largement dans le chapitre 4. Nous proposons ensuite une méthodologie d'évaluation de nos approches. Les performances de nos algorithmes sont analysées dans le chapitre 5 avant de conclure et de donner des perspectives à ce travail. Le présent rapport n'est qu'un résumé étendu du mémoire de thèse qui est détaillé en anglais à la suite de ce manuscrit, et dont l'organisation reste celle de cette synthèse.

Chapitre 1

Contexte clinique et problématique

1.1. Le cerveau humain

Anatomiquement intégré dans la boîte crânienne, le cerveau se situe au sommet de la hiérarchie fonctionnelle du système nerveux central. Appréhender la logique organisationnelle du tissu cérébral est aujourd'hui encore une question centrale de la neurobiologie moderne. Particulièrement développé chez l'homme, le cerveau est une structure vivante extrêmement complexe et non homogène, cette non-homogénéité s'exprimant entre autres par la juxtaposition de différents territoires auxquels peuvent être associées des fonctions plus ou moins bien définies. Le cerveau est de forme approximativement ovoïdale, de grand axe orienté antéro-postérieurement. Il comprend deux hémisphères séparés par un sillon médian, la scissure inter-hémisphérique, et reliés entre eux par des faisceaux tissulaires, les commissures inter-hémisphériques. La surface des hémisphères est recouverte d'une substance grise qui présente de nombreux plis dont les plus profonds s'appellent sillons ou scissures. Si l'on considère la face latérale d'un l'hémisphère cérébral, on distingue chez l'homme trois sillons principaux ou scissures, le sillon central ou scissure de Rolando, le sillon latéral ou scissure de Sylvius, le sillon occipital transverse ou scissure pariéto-occipitale. Les scissures délimitent ce que l'on appelle les lobes : le lobe frontal, le lobe pariétal, le lobe temporal, le lobe occipital. La surface des lobes est parcourue par des sillons moins profonds que les scissures qui délimitent de gros plis de substance grise appelés circonvolutions cérébrales ou gyri. La face médiale d'un hémisphère, quant à elle, n'est visible que par section des commissures inter-hémisphériques. Sur cette face médiale, on distingue une circonvolution corticale particulière, appelée circonvolution limbique ou gyrus cingulaire. La partie inférieure de la circonvolution limbique est enroulée sur elle-même. Elle est située contre un repli profond, appelé hippocampe. C'est une circonvolution inversée, repliée vers l'intérieur du cerveau.

Le cerveau contient deux types de substance : la substance blanche centrale et la substance grise, la première constituée de gaine de myéline (autour des axones assurant des connexions au-delà d'un voisinage immédiat), et la seconde de neurones et synapses. Située en périphérie (autour

de la substance blanche) dans l'encéphale, la substance grise, très plissée, constitue l'écorce cérébrale ou cortex et est formée de cellules nerveuses disposées en six couches. Autour des ventricules cérébraux, la substance grise est agglutinée en amas cellulaires, qui n'ont pas l'organisation laminaire des cortex. Ces amas sont appelés noyaux gris centraux. La substance blanche est un tissu du système nerveux central qui constitue la partie interne du cerveau et la partie superficielle de la moelle épinière. Elle occupe l'espace compris entre le cortex, les noyaux gris centraux et les ventricules.

1.2. L'épilepsie

L'épilepsie est l'un des désordres neurologiques les plus courants et n'a pas de frontière en termes d'âge, de race, de milieu social ou géographique. Elle touche les hommes et les femmes et peut débuter à n'importe quel âge mais est plus souvent diagnostiquée au moment de l'enfance, de l'adolescence ou à un âge avancé. Les raisons génétiques congénitales et développementales sont plus souvent associées à des patients jeunes alors que les tumeurs relèvent plus des adultes (au-delà de 40 ans). N'importe qui peut être sujet à des crises. Jusqu'à 5% de la population mondiale peut connaître une crise à un moment de sa vie mais une épilepsie n'est diagnostiquée qu'après la survenue d'au moins deux crises avérées. La prévalence de l'épilepsie est estimée à 0,5% de la population. Son incidence est de 50 nouveaux cas chaque année pour 100000 sujets.

Par le passé, l'épilepsie était associée à des expériences religieuses et faisait référence à des notions de démons. Ce n'est qu'au 19^{ème} siècle que la neurologie est apparue comme une discipline distincte de la psychiatrie et que le concept de désordre cérébral a été plus facilement accepté, plus spécialement en Europe et aux Etats-Unis. Le développement en 1920 de l'électroencéphalogramme suite aux travaux de Hans Berger en Allemagne a permis de mettre en évidence des décharges électriques au niveau cérébral différentes selon les types de crises et a aidé dans la localisation des décharges épileptiques ouvrant ainsi, en synergie avec les avancées de l'imagerie anatomique, la voie à une neurochirurgie ciblée.

Maladie neurologique traduisant un fonctionnement anormal, aigu et transitoire de l'activité électrique du cerveau, l'épilepsie se manifeste par des phénomènes paroxystiques survenant sur des plages temporelles relativement courtes (de quelques dizaines de secondes à plus d'une minute) appelées crises. Ces phénomènes sont classiquement à mettre en relation avec une activité anormale, hypersynchrone, d'une ou plusieurs populations de neurones impliquées au démarrage et durant le déploiement de la crise. Cette "décharge" épileptique critique est objectivée par l'enregistrement de l'ElectroEncéphaloGramme (EEG), qui peut être effectué en surface par des capteurs de potentiels disposés en maillage à la surface du crâne ou en profondeur au moyen de capteurs disposés sur des aiguilles insérées chirurgicalement (petit trou percé dans la calotte crânienne, dont le positionnement est contrôlé par imagerie). Ces crises répétitives, souvent imprévisibles, soudaines et brèves, peuvent prendre des formes très diverses. Compte tenu des multiples formes d'expression des crises et de la variété de leur évolution, on parle plutôt "des" épilepsies et non de l'épilepsie. L'observation directe des crises et, *a fortiori*, leur enregistrement EEG, font le plus souvent défaut en pratique courante. Le diagnostic repose essentiellement sur un interrogatoire du patient, et la description scrupuleuse et

précise du déroulement de la crise, ainsi que sur des enregistrements EEG, obtenus en absence de crise, appelés enregistrements intercritiques. Le récit du patient et/ou de son entourage permet d'apprécier l'existence de signes évocateurs de la maladie : mouvements convulsifs, perte de connaissance, chute, absences, relâchement des sphincters, automatismes...

Tout individu peut être concerné par une première crise d'épilepsie, sans que, dans la moitié des cas, il n'y en ait d'autres. Actuellement, un consensus se dégage pour considérer que connaître deux crises suffit pour être diagnostiqué comme "épileptique". Les examens morphologiques comme la tomodensitométrie et surtout l'imagerie par résonance magnétique nucléaire peuvent permettre un diagnostic étiologique précoce, sans nécessité d'hospitalisation. Cependant l'étiologie des épilepsies demeure inexpiquée dans plus de 50% des cas.

Dans 40% des cas, les épilepsies sont associées à une lésion cérébrale (malformation congénitale, encéphalite, séquelles d'une souffrance à la naissance, traumatisme crânien, accident vasculaire cérébral, tumeur, infections du système nerveux central, maladies neurologiques évolutives, anomalies des chromosomes, malformations cérébrales...). On les qualifie d'épilepsies symptomatiques. A côté de ces épilepsies, on trouve les épilepsies cryptogéniques, lorsqu'une cause est suspectée, mais ne peut être prouvée par les méthodes diagnostiques actuelles. Leur nombre est en constante diminution du fait de l'évolution permanente des techniques d'exploration. Entre 5 et 10% des épilepsies sont d'origine génétique. Ce sont les épilepsies idiopathiques. Elles concernent des sujets le plus souvent sans lésion cérébrale. La prédisposition génétique n'induit pas que l'épilepsie soit une maladie héréditaire et sa transmission est très complexe, la même anomalie pouvant avoir des conséquences différentes selon les personnes. Les cas d'épilepsie provoqués par une maladie, transmise elle-même génétiquement, sont à considérer différemment. Ce sont essentiellement des épilepsies datant de l'enfance ou de l'adolescence, n'ayant pas guéri, mais le plus souvent en rémission durable sous traitement.

Indépendamment de leur étiologie, les crises peuvent être classées en deux grands types : les crises généralisées, qui impliquent l'ensemble du cerveau et les crises partielles, limitées à une région précise du cerveau. Il arrive qu'une crise, d'abord partielle, diffuse à l'ensemble du cerveau et devienne ainsi généralisée. Les épilepsies, qu'elles soient symptomatiques ou cryptogéniques, sont le plus souvent partielles et représentent près de 80% des épilepsies de l'adulte. Quant aux épilepsies idiopathiques (pas de cause décelée) de l'adulte, elles sont toujours généralisées.

Dans les crises généralisées, on distingue :

- les crises tonico-cloniques : ce sont les plus connues, les plus impressionnantes mais pas les plus fréquentes. Elles se manifestent par une perte de connaissance avec chute, mouvements convulsifs, risque de morsure de la langue, ...
- les absences : elles se manifestent par une brève rupture de contact (quelques secondes), se traduisant par une fixité du regard ; elles sont parfois accompagnées de mâchonnements, ou de gestes involontaires et inadaptés appelés automatismes. La

personne perd le contact avec son environnement mais conserve son tonus musculaire. Les absences peuvent se répéter de manière fréquente au cours de la journée.

Dans les crises partielles, qui n'affectent que certaines parties du corps, on distingue :

- les crises partielles simples, qui durent généralement quelques minutes. Les symptômes dépendent de la zone du cerveau atteinte. Elles peuvent se traduire par des troubles moteurs, des troubles sensoriels et sensitifs, des troubles de la mémoire ou de la conscience. Certaines de ces crises partielles peuvent évoluer vers une crise généralisée tonico-clonique ;
- les crises partielles complexes, durant lesquelles l'individu est dans un état de conscience altérée. Il ne répond pas aux stimulations et son regard est fixe. Il peut avoir des automatismes, c'est-à-dire qu'il manifeste des gestes répétitifs involontaires comme tirer sur ses vêtements ou claquer des dents. Une fois la crise finie, il ne se souvient pas du tout ou très peu de ce qui s'est passé.

1.3. Vers quelles solutions ?

L'épilepsie reste difficile à traiter dans bien des cas, et des progrès tant sur le versant diagnostique que thérapeutique ne sont obtenus qu'au prix d'efforts importants, incluant des actions concertées alliant technologues et médecins.

1.3.1. Les avancées scientifiques et médicales

La recherche en épilepsie relève de deux axes : la recherche fondamentale et la recherche clinique. La recherche fondamentale s'attache à comprendre les mécanismes sous-jacents au développement de la maladie, sa cause, ses différentes manifestations, ses conséquences sur les fonctions cérébrales. Mieux comprendre les processus cérébraux responsables des crises peut conduire à de nouvelles approches de prévention et de traitement (nouvelles drogues anti-épileptiques, identification de gènes...). Quant à la recherche clinique, elle est principalement concernée par l'application de nouvelles technologies de diagnostic et de thérapie. Aujourd'hui, les technologies disponibles permettent aux scientifiques de mieux appréhender le cerveau humain : l'Imagerie par Résonance Magnétique (IRM), la Tomographie par Emission de Positons (TEP), la StéréoElectroEncéphaloGraphie (SEEG), la MagnétoEncéphaloGraphie (MEG)...

Pour traiter l'épilepsie, on a recours à deux types de thérapie : la pharmacothérapie et la chirurgie. Si les traitements sont avant tout médicamenteux (les principaux traitements incluent encore aujourd'hui la phénytoïne et le phénobarbital introduits au début du siècle dernier), il n'existe pas d'anti-épileptique spécifique d'une forme d'épilepsie donnée. Seule l'expérience du neurologue peut guider le choix du traitement. La prise régulière et quotidienne du traitement est le seul garant d'efficacité sur les crises.

Trois réponses au traitement sont habituellement observées :

- les crises disparaissent assez rapidement après la mise en route du traitement ;

- les crises disparaissent mais les risques de rechute restent importants à l'arrêt du traitement ;
- dans 10 à 20 % des cas, les crises persistent malgré toutes les tentatives de traitement médicamenteux. On parle d'épilepsie pharmaco-résistante.

S'il s'agit de crises partielles et que le traitement n'apporte pas de résultats tangibles, il est alors possible de procéder à une intervention chirurgicale. Celle-ci consiste à pratiquer une incision dans la partie du cerveau liée aux crises, ou encore à retirer complètement cette partie, si le risque de provoquer un déficit neurologique peut être exclu. L'évaluation qui précède la chirurgie peut durer plusieurs mois, comprenant de multiples examens (tests d'imagerie, électroencéphalogramme, évaluations neuro-psychologiques). Cette chirurgie est proposée dans 5 à 6% des épilepsies pharmaco-résistantes. A titre indicatif, on considère qu'en France 40 000 personnes souffrent d'un handicap grave dû à une épilepsie partielle pharmaco-résistante.

1.3.2. Problématique de l'étude

Ce travail de thèse s'inscrit dans la problématique de la définition (détermination), pour un patient donné, de la (des) région(s) du cerveau candidate(s) à une excision de sorte que les crises soient supprimées ou fortement atténuées, tout en limitant impérativement les déficits post-chirurgicaux (sensitifs, moteurs ou cognitifs) induits par l'intervention. Il s'agit donc de circonscrire, à partir d'observations, la zone épileptogène, responsable des crises. En fait cette zone épileptogène peut correspondre à un réseau d'ensembles neuronaux, distribués à travers des structures parfois éloignées, et supportant l'initiation et la propagation des activités épileptiques. De par ce caractère distribué, la localisation et la caractérisation des réseaux épileptogènes restent un problème difficile. Elles sont cependant fondamentales car elles seules peuvent aboutir à la délimitation d'une région corticale minimale à exciser pour éradiquer la survenue des crises.

Dans ce contexte, on comprend l'enjeu que représente l'analyse de signaux électroencéphalographiques intracérébraux enregistrés chez les patients qui vise à (i) identifier les structures cérébrales impliquées dans les différentes phases d'une crise (repérer les nœuds d'un graphe de connectivité), (ii) quantifier l'information marginale portée par les différents signaux (attributs pour chaque nœud) et (iii) caractériser les relations inter-structures au cours du temps via l'analyse conjointe de ces signaux (liens statistiques et détection d'influences causales). Nos travaux ont pour objectif de contribuer au développement de ce troisième point, *i.e.* à la quantification de flux d'informations entre structures en détectant/analysant des relations au sein d'un ensemble de signaux acquis sur différents capteurs, afin de mieux comprendre l'organisation de la crise en termes de propagation. L'aspect évolutif, non stationnaire, de cette organisation est pris en compte principalement en se donnant une contrainte de fenêtre temporelle d'analyse courte (de l'ordre de quelques secondes).

Pour ce faire, notre étude portera *in fine* sur l'analyse de signaux caractéristiques d'une activité épileptique, soit simulée par un modèle physiologique, soit obtenue *ex vivo*. Toutefois, les approches développées seront d'abord appliquées à des signaux simulés au moyen de modèles autorégressifs

vectoriels linéaires ou non avant donc d'être testées sur des modèles physiologiques et enfin sur des signaux réels obtenus sur un modèle animal. Ces derniers sont recueillis lors de crises d'épilepsie induites chez le cochon d'inde, l'activité recueillie présentant des points communs avec celle relevée chez l'homme durant l'initiation de la crise. Le modèle animal à l'origine de ces données, développé au Département de Neurophysiologie Expérimentale de l'Institut Neurologique (Carlo Besta) de Milan (par le Professeur M. de Curtis) consiste à induire des crises d'épilepsie par injection de bicuculline sur des cerveaux de cochons d'inde isolés (mais maintenus en vie par perfusion durant une dizaine d'heures), et à enregistrer l'activité électrique résultante. Pour cela sont utilisées des électrodes laminaires procurant seize voies d'enregistrement, introduites perpendiculairement au cortex entorhinal. Ce dernier joue *a priori* un rôle déterminant dans l'initiation des crises temporales, comme le confirment de nombreuses études effectuées chez l'homme.

Actuellement, il semble se dégager cinq grandes phases dans l'organisation de ce type de crise, une phase intercritique, une activité précritique, une activité tonique à bande étroite et à fréquence élevée (FOA : Fast Onset Activity), une activité clonique qui se divise en un mélange de bouffées rapprochées et d'intervalles inter-bouffées plus organisés, une activité post-critique normale se mettant en place en fin de crise. Les interrogations portent sur l'organisation en réseau des structures cérébrales impliquées lors des différentes phases d'une crise, et plus particulièrement durant les décharges rapides apparaissant lors de sa mise en place (FOA).

Pour répondre à ces questions, au-delà de la connectivité fonctionnelle qui renvoie à la notion de couplage statistique entre signaux obtenus dans différentes régions impliquées, il s'agit d'inférer la connectivité effective, *i.e.* d'établir des graphes traduisant des flux d'informations entre populations impliquées, à partir de ces mêmes signaux. Pour cela, différentes approches foisonnent dans la littérature, avec une mise en exergue fréquente des limites des méthodes dites linéaires contrecarrées par le potentiel de méthodes non linéaires non paramétriques, principalement basées sur des fonctionnelles exprimables sous forme d'entropies marginales et conjointes, comme l'information mutuelle, si l'on vise une connectivité fonctionnelle, et l'information de transfert, si l'on s'intéresse à une approche effective. Il s'avère d'ores et déjà difficile *a priori* de s'orienter (i) pour choisir un "bon" indicateur de connectivité au sens de ses performances asymptotiques mais également correctement estimable pour des échantillons de taille limitée, (ii) pour valider les performances des méthodes tant en simulation, dans la mesure où ces dernières sont appliquées sur des modèles (même si ceux-ci sont tout à fait réalistes, ils n'en restent pas moins des "modèles"), qu'en situation réelle (pour laquelle la référence "terrain" n'est pas disponible).

Notre contribution tente d'apporter des repères en investissant des approches linéaires et d'autres, non linéaires, dont certaines sont originales, pour les appliquer à des signaux réels, ou du moins simulés de manière physiologiquement argumentée. Des éléments de réponses seront ainsi donnés au problème du choix des indicateurs pour découvrir (i) quelles sont les structures impliquées dans des activités rapides, (ii) quelle est la connectivité fonctionnelle/effective entre les régions observées et (iii) quelles structures jouent un rôle "leader" dans l'initiation des crises et leur propagation.

Chapitre 2

Connectivité et modèles

Notre objectif étant de contribuer aux méthodologies de mesure de la connectivité cérébrale à des fins de diagnostic de l'épilepsie, deux points sont particulièrement discutés dans ce chapitre : la connectivité et ses différents attributs, les modèles de l'épilepsie eu égard à ses différentes échelles d'observation. Cette caractérisation de la connectivité cérébrale est aujourd'hui encore largement discutée et en dominer les différents enjeux nous est apparu indispensable avant de privilégier un axe de recherche. De même, une connaissance suffisamment fine des modèles computationnels de la littérature permettra de mieux comprendre et de justifier les choix qui seront faits par la suite pour tester nos approches en simulation.

2.1. Connectivité cérébrale

Au fil des années, on note un intérêt croissant dans la compréhension des fonctions cérébrales normales et pathologiques, et, dans ce contexte, l'identification de variations dans l'activation des aires cérébrales et de leurs interactions est riche d'enseignements. Comprendre et modéliser les fonctions du cerveau nécessite donc non seulement d'identifier correctement les régions activées mais aussi de déceler les interactions fonctionnelles parmi les ensembles neuronaux éventuellement distribués au sein du cortex. Ces concepts font référence aux principes de ségrégation fonctionnelle (activation d'ensembles neuronaux dans des régions spécialisées) et d'intégration (activation coordonnée de groupes de neurones largement distribués dans des régions potentiellement distantes du cortex) [Friston 2009].

L'intégration d'aires cérébrales peut se mesurer par l'évaluation de la connectivité cérébrale qui couvre trois notions fondamentales : la connectivité anatomique ou structurelle, qui a trait à l'organisation anatomique des liaisons corticales, la connectivité fonctionnelle qui se réfère à la notion de dépendance statistique et la connectivité effective qui sous-tend la notion d'interactions causales entre des unités distinctes au sein d'un même système nerveux. Les unités sont prises ici au sens élargi et peuvent tant représenter des neurones individuels que des populations neuronales ou encore des aires cérébrales. Le "pattern" de connectivité peut ainsi être représenté par des connexions

structurelles que constituent les axones et leurs synapses ou (à une échelle plus grossière) les réseaux de fibres, ou au moyen de mesures de relations statistiques comme l'intercorrélation, la cohérence... ou encore de relations causales comme les mesures de flux d'information. L'activité neuronale, et par extension les codes neuronaux, sont contraints par cette connectivité dont la connaissance peut permettre d'élucider les processus de traitement de l'information au niveau des neurones et des réseaux associés [Edeline 1999, Sakalis 2011, Sporns 2004, Sporns 2010, Treserras 2008].

Dans cette partie, les différents aspects de l'organisation cérébrale seront d'abord discutés avant de synthétiser les techniques actuelles qui explorent ces concepts de connectivité.

2.1.1. Généralités

2.1.1.1. Modes de connectivité

Comme évoqué précédemment la connectivité intègre les notions fondamentales de connectivité structurelle, fonctionnelle et effective [Friston 1994, Sporns 2004, Sporns 2005, Sporns 2007].

La connectivité anatomique correspond à un réseau de connexions physiques ou structurelles (synaptiques) reliant des ensembles de neurones ainsi que leurs caractéristiques structurelles associées résumées sous forme de paramètres comme l'efficacité de la connexion synaptique. Sur une échelle temporelle courte, *i.e.* de quelques secondes à quelques minutes, la connectivité anatomique est stable, alors qu'au-delà, *i.e.* de quelques heures à quelques jours, elle peut être reconfigurée en raison de changements morphologiques significatifs ou de phénomènes de plasticité.

La connectivité fonctionnelle se rapporte à un concept statistique. Elle est le reflet de modifications de dépendance (ou d'indépendance) statistique entre les activités d'unités neuronales distribuées et éventuellement distantes spatialement. Cette dépendance statistique peut être estimée par des mesures connues de corrélation, covariance ou cohérence spectrale. Le plus souvent elle l'est entre tous les éléments du système étudié, quelles que soient les connexions directes ou indirectes qui existent entre ces éléments. A l'inverse de la connectivité structurelle, la connectivité fonctionnelle peut être hautement non stationnaire. Ainsi, les "patterns" statistiques représentatifs de cette connectivité fluctuent à des échelles temporelles variables, pouvant atteindre la dizaine de millisecondes, voire moins. Cette connectivité ne fait référence à aucune forme de directivité ni à un modèle structurel sous-jacent.

Enfin, la connectivité effective peut être considérée comme cherchant à synthétiser connectivités structurelle et fonctionnelle, dans le sens où elle décrit des réseaux d'influence d'un élément neuronal sur un autre. Elle tient compte du fait que les processus neuronaux ne violent pas le principe de causalité introduit en physique. En principe, les effets de causalité peuvent être inférés par l'introduction de perturbations sur le système ou, puisque les causes précèdent les effets, par l'analyse de séries temporelles comme nous allons le voir ultérieurement. Pour détecter cette connectivité effective, les techniques développées peuvent ou non requérir la spécification d'un modèle paramétrique.

2.1.1.2. *Le formalisme des graphes*

La connectivité peut être étudiée sous l'angle de l'analyse de réseaux. Dans ce cadre, différents modes de représentation de connectivité, comme les graphes ou les matrices, s'avèrent intéressants puisqu'ils peuvent s'appliquer aux différents types de connectivité et ce à n'importe quelle échelle. Formellement, un graphe correspond à la donnée d'un ensemble de sommets ("nœuds") et d'une application affectant à chaque paire de sommets une valeur binaire pour signifier l'existence ou non d'un lien, plus éventuellement une valeur numérique (graphe pondéré) pouvant s'interpréter différemment suivant les applications. Les nœuds correspondent aux neurones, dans le cas de connectivité à l'échelle microscopique, et aux régions cérébrales à des échelles plus macroscopiques. Ils peuvent interagir de manière directe ou indirecte via des chemins de propagation composés d'arêtes (dans le cas d'un graphe non orienté) ou d'arcs (dans le cas d'un graphe orienté). Ces arêtes ou "liaisons" correspondent aux synapses dans le cas microscopique ou à des réseaux de fibres ou de voies de propagation plus élargis à un niveau plus macroscopique. Les graphes peuvent être représentés par des matrices de connexion dont les éléments traduisent la présence ou l'absence d'arête entre paires de sommets. Chaque connectivité a sa propre représentation. Par exemple, dans le cas de connectivité anatomique, on traite de matrices relativement creuses dans la mesure où beaucoup de paires de nœuds ne sont pas reliées physiquement soit par une fibre soit par un faisceau de fibres. Dans sa forme la plus simple, le graphe est binaire, ses éléments indiquant la présence ou l'absence de connexions, mais il peut également s'agir d'un graphe pondéré dans lequel les éléments constitutifs sont pondérés par des coefficients traduisant la densité ou l'efficacité des connexions. Dans le cas de la connectivité fonctionnelle (au sens classique), la matrice est symétrique, et chaque élément code la dépendance statistique entre deux composants du système (neurones, sites d'enregistrements...). De telles matrices peuvent être seuillées pour produire des graphes non orientés binaires, le seuil permettant d'inclure dans le graphe résultant uniquement les liens dépassant un certain niveau d'intensité. Puisqu'une valeur significative d'efficacité fonctionnelle peut être mesurée entre des sommets indirectement liés, un graphe de connectivité fonctionnelle comprend généralement plus d'arêtes que le graphe structurel correspondant lorsque celui-ci est connu. Ainsi, quand deux sites neuronaux sont reliés par plus d'un chemin, certains ne correspondent pas à un chemin physiquement existant. Généralement, l'importance des relations indirectes dépend de la longueur du chemin définie sur le graphe structurel. Finalement, concernant la connectivité effective, contrairement à la précédente, elle conduit à une matrice non symétrique puisque certains liens peuvent être unidirectionnels ou bidirectionnels et dans ce dernier cas d'importances non identiques. Les liens du graphe sont donc orientés, et, là encore, un seuillage conduit à la définition de graphes orientés binaires. Différentes mesures, correspondant à différentes fonctionnelles (si le graphe est assimilé à une fonction sur les paires de nœuds), ont été introduites pour quantifier globalement certaines caractéristiques relatives par exemple à un nombre moyen de liens directs ou indirects. Notons que, sans connaissance de liens anatomiques, il est impossible, à partir des mesures de connectivités fonctionnelle et effective encodées dans le graphe, de savoir si une arête donnée correspond ou non à un ensemble effectif de fibres. En résumé, il est possible de détecter une arête fonctionnelle sans qu'il y ait d'arête structurelle et, inversement, de reconnaître un lien structurel sans lien statistique significatif entre deux nœuds.

2.1.2. Techniques d'observation et approches computationnelles

2.1.2.1. Connectivité structurelle

La connectivité structurelle est intrinsèquement difficile à définir rigoureusement car, à l'échelle microscopique des neurones, de nouvelles connexions synaptiques apparaissent alors que d'autres disparaissent, cette organisation dynamique pouvant être liée à une fonction exécutive [Ooyen 2001]. Les niveaux d'analyse de la connectivité structurelle vont des connexions synaptiques individuelles, à l'échelle microscopique, jusqu'aux connexions entre régions corticales à l'échelle macroscopique en passant par l'analyse de populations de neurones (arrangement en colonnes corticales).

Quelle que soit l'échelle d'observation (microscopique, mésoscopique ou macroscopique), les connexions anatomiques sont caractérisées par deux attributs : la spécificité et la variabilité. La spécificité provient de l'organisation particulière des connexions synaptiques individuelles en fonction des types (morphologiques et physiologiques) de neurones et de la connectivité à plus ou moins longue portée entre les structures neuronales (que ce soit au niveau des noyaux cellulaires que des régions corticales explorées). La variabilité se mesure soit entre structures correspondantes au sein d'une même espèce soit au sein d'un même individu au cours du temps prenant en compte les phénomènes de développement (au sens de croissance) et de plasticité. Il est probable que cette variabilité anatomique soit la plus grande cause de variabilité fonctionnelle.

Les premières investigations de connectivité structurelle étaient limitées aux techniques de dissection post-mortem, qui permettaient seulement une description assez grossière de la localisation et des orientations des principales fibres. Au cours de la dernière décennie, avec l'avènement de l'IRM (imagerie par résonance magnétique) de diffusion, les nouvelles techniques telles que l'imagerie par tenseur de diffusion (DTI : Diffusion Tensor Imaging) ou encore l'imagerie du spectre de diffusion (DSI : Diffusion Spectrum Imaging) (plus sensible aux hétérogénéités intra-voxel dues au croisement de fibres et permettant un meilleur rendu des trajectoires des axones) ont permis des avancées dans l'examen de la connectivité structurelle. Dans ce panorama, le "phénomène du petit monde" ("small-world phenomenon") a été très largement recherché et bien souvent rencontré. Au niveau structurel, il apparaît que le cortex est constitué de "clusters", c'est-à-dire de petits groupes élémentaires présentant des propriétés similaires, globalement interconnectés, mis en évidence par ces techniques d'imagerie. Globalement, ces dernières révèlent un réseau cortical fortement fragmenté en petits groupes, dans lequel existent principalement des voies d'accès entre des aires spatialement proches. L'analyse des contributions structurelles des aires individuelles permet l'identification et la classification de pôles hautement connectés dans des régions centrales du cerveau. Ces régions incluent des aires du cortex pariétal et préfrontal, ce qui explique en partie leur activation lors de tâches cognitives. Les réseaux structurels du cortex humain ne sont pas encore pleinement caractérisés à ce jour mais l'utilisation de techniques d'imagerie de diffusion non invasives ouvre le champ à d'importantes avancées. Des techniques alliant différentes modalités apparaissent aussi très prometteuses comme celle alliant stimulation magnétique transcrânienne et tomographie par émission de positons [Paus 1997] qui a déjà permis de mieux préciser *in vivo* la cartographie du cerveau humain.

2.1.2.2. *Connectivités fonctionnelle et effective*

2.1.2.2.1. **Modalités instrumentales**

Si l'analyse des réseaux structurels aide à la compréhension de l'architecture des connexions entre régions, la considération de réseaux fonctionnels est indispensable pour comprendre comment cette architecture intervient dans les dynamiques neurophysiologiques.

La connectivité fonctionnelle est définie comme la corrélation temporelle (en termes de dépendance statistiquement significative entre régions) d'activités d'ensembles neuronaux. Moults signaux neurophysiologiques peuvent être caractérisés par des techniques associées et, dans ce contexte, des approches basées sur l'imagerie fonctionnelle et sur l'électrophysiologie s'avèrent très complémentaires, l'une bénéficiant d'une bonne résolution spatiale, la seconde d'une bonne résolution temporelle. Là encore, l'investigation de techniques d'exploration multi-modales vient enrichir ce concept de connectivité fonctionnelle permettant de comprendre non seulement comment des régions s'activent mais surtout comment elles communiquent entre elles au cours d'une tâche cognitive par exemple ouvrant la voie vers la connectivité effective. Ainsi, par exemple, dans [McIntosh 1994], l'analyse de données cérébrales passe par des covariances calculées sur des données multi-modales. La technique, appelée SEM (Structural Equation Modeling) et sur laquelle on reviendra ultérieurement, se réfère à la mise en équation linéaire d'une modélisation structurelle basée sur des connaissances anatomiques *a priori*. Des covariances d'activité sont utilisées pour identifier des coefficients, chacun représentant la force d'influence d'une structure sur une autre. De même la causalité de Granger a été mise à profit dans l'examen de signaux LFP (Local Field Potentials) recueillis lors de tâches comportementales et cognitives. La combinaison d'approches, alliant par exemple la stimulation magnétique transcrânienne avec d'autres fonctionnalités, a déjà largement fait ses preuves pour comprendre les liens entre des régions localement distantes du cerveau [Bestmann 2008, Ruff 2009], mais aussi pour évaluer la connectivité cérébrale dans des situations pathologiques indépendamment des capacités cognitives, sensorielles ou motrices des patients [Ferrerri 2011, Hallett 2000, Paus 1997], ou encore pour mieux caractériser les mécanismes d'inhibition et de facilitation entre le cortex et la voie cortico-spinale [Massimini 2005].

2.1.2.2.2. **Traitement des données dans la connectivité fonctionnelle**

Concernant l'évaluation de la connectivité fonctionnelle, deux types d'approches se dégagent : les techniques linéaires et les techniques non linéaires.

2.1.2.2.2.1. *Techniques linéaires*

Ce sont dans les années 50 que l'on voit apparaître les premières mesures de connectivité linéaire cérébrale par l'introduction de mesures d'intercorrélation sur des paires de signaux EEG et de corrélation partielle sur des signaux trivariés. Pour mesurer cette connectivité linéaire dans le domaine fréquentiel, sont introduites les notions de cohérence (cohérence ordinaire entre deux signaux ou partielle sur des signaux multivariés). La cohérence mesure des corrélations spatiales (calculées entre signaux indexés, chacun, par une position de capteur) et dans différentes bandes de fréquences. Cette quantité complexe (décrite par une amplitude et une phase) est sensible aux changements de relations statistiques, fréquence par fréquence, entre la paire (module, phase) du premier signal et

celle du second. S'il existe une relation linéaire et invariante dans le temps entre les signaux considérés, le module de la cohérence est unité.

2.1.2.2.2. *Techniques non linéaires*

Non nécessairement conçues pour suppléer les techniques linéaires, les techniques non linéaires peuvent apporter des informations complémentaires. Issue de l'analyse des systèmes dynamiques non linéaires et souvent liée à la notion de chaos, une première famille de méthodes a été développée [Lorenz 1963]. L'analyse de séries temporelles non linéaires a été historiquement motivée par le fait que de nombreux processus neuronaux présentent des caractéristiques non linéaires.

Au début des années 80, le concept de synchronisation, ou de coordination de systèmes, a émergé et connu une série de déclinaisons, dont la synchronisation de phase qui prend son sens si l'on s'intéresse à la simultanéité ou au délai de réalisation (fraction de période) de deux événements cycliques. On retrouve effectivement cette notion dans l'étude de systèmes dynamiques chaotiques, et plus particulièrement de leurs attracteurs (un attracteur correspond à un sous-ensemble de l'espace d'état sur lequel toute trajectoire d'état vient se confiner asymptotiquement, quelle que soit son origine dans un ensemble d'attraction incluant l'attracteur). La synchronisation de phase correspond à la situation où deux systèmes oscillants chaotiques couplés évoluent avec une même période et simultanément alors que les enveloppes de leurs oscillations peuvent être décorréliées. Par extension, la synchronisation généralisée correspond à l'existence d'une relation fonctionnelle instantanée entre les trajectoires d'état respectives des deux systèmes. La synchronisation peut donc être vue comme un ajustement de rythmes d'objets oscillants. Dans le contexte des neurosciences, ces deux types de synchronisation ont fait l'objet de nombreuses études qui ont cherché à établir des indices pour quantifier les similarités de trajectoire des signaux observés et prouvé leur apport dans l'examen et l'explication de mécanismes inhérents à certaines pathologies, comme la genèse de phénomènes épileptiques. Une seconde famille de méthodes est issue de la théorie de l'information de Shannon, la méthode la plus représentative dans ce domaine étant l'information mutuelle qui mesure la dépendance statistique entre deux variables aléatoires. Celle-ci est nulle si, et seulement si, les variables sont indépendantes, et croît avec leur dépendance. En d'autres termes, elle mesure la quantité d'information apportée en moyenne par la réalisation d'une variable sur la réalisation de la seconde variable. Cette quantité, positive ou nulle, est symétrique en ses deux arguments. L'absence d'information mutuelle peut s'exprimer à partir de l'entropie, et revient à dire que l'entropie conjointe des variables considérées est égale à la somme des entropies de chacune des variables. Au-delà de la situation bivariée, l'information mutuelle partielle représente la quantité d'information commune à (ou partagée par) deux variables aléatoires, conditionnellement à la connaissance d'une troisième variable (et toujours en moyenne). Naturellement, si cette troisième variable est indépendante des deux premières, l'information mutuelle partielle est égale à l'information mutuelle.

2.1.2.2.3. Traitement des données dans la connectivité effective

La connectivité effective est un concept beaucoup plus récent défini comme l'influence - directe ou non - qu'exerce un système sur un autre, cette influence pouvant être réciproque [Horwitz 2003]. De manière naturelle, la connectivité effective s'interprète comme une synchronisation et une interaction d'ensembles neuronaux s'exprimant dans des régions proches ou distantes, soit pour accomplir des fonctions perceptives, motrices ou cognitives, soit comme une conséquence d'états pathologiques comme ceux causés par l'épilepsie. Dans les processus cognitifs, les fonctions mises en œuvre (anticipation de l'occurrence d'un stimulus, attention portée à ce stimulus, actions qu'il va potentiellement provoquer) impliquent des interactions complexes. Un tel processus d'interactions s'accompagne de couplages bidirectionnels et/ou unidirectionnels. On peut relier les premiers à une synchronisation mutuelle, où deux systèmes (ou davantage) ajustent leurs rythmes l'un à l'autre. Une synchronisation mutuelle entre deux systèmes A et B peut s'interpréter comme une double causalité, où l'évolution de l'état d'un système est explicitement fonction de son propre état (actuel) et de celui du second système, et les seconds à une interaction causale dirigée entre le système initiateur ou système maître ("driving system") et sa réponse ou système esclave ("driven system").

Aujourd'hui, la conceptualisation de cette connectivité est encore débattue : on peut la formaliser en considérant donc des liens causaux, *i.e.* à partir d'une combinaison des connectivités structurelle et fonctionnelle, ce qui implique l'introduction d'un modèle de connaissance, ou l'estimer directement à partir des observations, *i.e.* pilotée par les données [Sakkalis 2011], sans introduire explicitement de connaissances physiologiques. Ces techniques sont résumées ci-après.

2.1.2.2.3.1. Techniques basées sur un modèle

L'évidence neurobiologique et les théories plausibles qu'elle génère sont à l'origine de modèles théoriques permettant de décrire les interactions entre structures cérébrales mais aussi la manière dont celles-ci s'influencent mutuellement. Ce concept de connectivité effective basée sur les modèles voit aujourd'hui la compétition de différents modèles neurobiologiques avec leurs propres hypothèses.

Proposée par McIntosh et Gonzalez-Lima [McIntosh 1994], la modélisation par équation structurelle (SEM) de la covariance assigne des variables de connectivité effective aux voies anatomiques les plus pertinentes (ou considérées comme telles *a priori*). Ces variables peuvent alors être estimées par maximisation de la vraisemblance des covariances calculées à partir de données enregistrées durant une tâche répétée. Une généralisation de cette approche est la modélisation causale dynamique (DCM : Dynamic Causal Modeling) [Friston 2003], dont l'idée maîtresse est de considérer le cerveau comme un système déterministe dynamique non linéaire soumis à différentes entrées et régi par trois types de paramètres, les premiers représentatifs de l'influence des entrées externes sur les états du système, les seconds représentatifs du couplage intrinsèque entre états, les troisièmes intervenant sur les entrées pour moduler le couplage. Les seconds correspondent à la connectivité effective, les derniers reflétant les changements dans la connectivité induits par les entrées. Cette approche se distingue des approches conventionnelles, dans lesquelles on suppose que les réponses observées sont pilotées par un bruit intrinsèque (ou endogène). Ici, le modèle suppose que les réponses sont pilotées par des modifications dans les entrées externes. De manière

standard, les modèles autorégressifs multivariés supposent un système excité par des innovations stochastiques et des interactions linéaires liant le signal vectoriel à l'instant présent et son passé. Quant à la modélisation par équation structurelle, elle suppose des interactions linéaires et instantanées. En résumé, l'approche DCM se singularise en s'adaptant aux caractéristiques dynamiques et non linéaires des interactions neuronales, mais aussi en traitant le problème du suivi (estimation) des paramètres de connectivité expliquant les données expérimentales. Cette technique a d'abord été introduite en IRM fonctionnelle avant d'être appliquée à l'EEG et à la MEG.

Une autre approche, se situant à un niveau différent, et développée pour caractériser la capacité d'intégration de l'information des réseaux cérébraux est proposée dans [Tononi 2003]. Elle permet d'identifier les régions cérébrales hautement interactives ainsi que leurs influences respectives à partir de la notion d'information effective. L'information effective entre deux éléments A et B, obtenus en partitionnant un système S, est égale à l'information mutuelle entre A et B lorsque l'on substitue aux variables constitutives de A (informations nodales) des variables indépendantes de mêmes lois marginales (technique comparable à la génération de "surrogate data" en traitement du signal), maximisant ainsi l'entropie de A. Cette maximisation permet de connaître le degré maximum potentiel d'influence de A sur B, inhérent à la seule connectivité. Une procédure symétrique identique est appliquée pour évaluer l'influence de B sur A. Cette démarche peut être répétée sur d'autres partitions de S, pour une évaluation moyenne (globale) de la connectivité au sein de S. Cette méthode est intéressante conceptuellement parlant, mais nécessite, dans sa forme actuelle, de manipuler un modèle gaussien dont le graphe est connu *a priori* en supprimant certains liens et en recalculant à chaque fois certaines entropies théoriques. Ceci passe par le calcul de bornes supérieures d'indicateurs de causalité entre sous-systèmes, sur la base d'un modèle structurel préexistant.

2.1.2.2.3.2. *Techniques pilotées par les données*

Contrastant avec les techniques précédentes, les techniques fondées sur les données ne suppose pas de modèle spécifique ou de connaissance *a priori* sur les relations spatiales ou temporelles. La technique la plus connue est la causalité de Wiener-Granger (WGC : Wiener-Granger Causality) [Granger 1969, Wiener 1956]. En 1956, Norbert Wiener a été le premier à reconnaître l'importance d'un ordonnancement temporel dans l'inférence de relations causales. Ainsi, au sens de Wiener, il existe un lien de causalité d'un signal X vers un signal Y à l'instant t si la connaissance du passé de X , ajoutée à celle du passé de Y améliore la variance d'erreur de prédiction de Y_t utilisant la seule connaissance du passé de Y . Ainsi, si une variable peut être prédite par ajout de l'information passée d'une seconde variable mieux que par sa seule information passée, cette seconde variable peut être considérée comme "causale" vis-à-vis de la première. Cette définition très générale a donné lieu à la mise au point de multiples outils mathématiques. Ainsi, en 1982, Geweke [Geweke 1982] proposa une mesure de causalité basée sur la représentation spectrale d'un processus ARX multivarié (FGC : Frequency Geweke Causality) utilisée par exemple pour mettre en évidence des interactions fonctionnelles entre les aires corticales visuelles du chat. En temps comme en fréquence, ces concepts se sont généralisés au cas multivarié donnant naissance à diverses mesures. Plus récemment, celles-ci se sont étendues au cas non linéaire. Préalablement au travail de

Geweke, une méthode spectrale, dénommée "cohérence dirigée" (DCOH : Directed COHerence) a été proposée par Saito et Harashima [Saito 1981]. Celle-ci prend en compte une source additionnelle de bruit, modélisant la part non observée de l'activité cérébrale afin de tenir compte de possibles sources externes influençant les signaux observés. L'extension de ce formalisme à un plus grand nombre de voies a fait l'objet des travaux de Kamiński et Blinowska [Kamiński 1991]. Ils ont ainsi introduit la notion de fonction de transfert dirigée (DTF : Directed Transfer Function) dans le cas multidimensionnel pour déterminer les directions de propagation de l'information entre deux voies en tenant compte de l'ensemble des voies considérées. Cette méthode a été largement appliquée pour l'étude de signaux EEG enregistrés lors de différentes phases du sommeil, la localisation de foyers épileptiques et l'épileptogénèse. Dans le même esprit, la cohérence dirigée partielle (PDC : Partial Directed Coherence) a été préconisée par Baccalá et Sameshima [Baccalá 2001] pour fournir une image plus précise de connectivité dans le domaine fréquentiel que celle donnée par la cohérence dirigée, en particulier pour l'analyse simultanée de plus de deux observations. Cette fonction est équivalente à la DTF dans le cas bivarié mais, dans le cas multivarié, la PDC est capable de détecter non seulement les voies directes de propagation mais aussi les voies indirectes qui relient deux aires corticales interactives. Plus récemment, en parallèle de ces approches exploitant majoritairement l'amplitude de fonctions de transfert, est apparue une mesure, l'indice de pente de phase (PSI : Phase Slope Index) [Nolte 2008] qui exploite la monotonie de la phase entre deux signaux pour identifier des relations de causalité dans les situations de propagation de flux unidirectionnel. Le principe de cette technique est d'identifier les retards relatifs entre composantes spectrales des signaux dans les bandes de fréquences où la cohérence est significative. Si les méthodes précédentes ont reçu un large écho dans la littérature sur la connectivité effective, elles peuvent être mises en défaut sur des systèmes neurophysiologiques plus complexes, justifiant l'intérêt que des chercheurs ont préféré porter à des approches relevant de la théorie de l'information. Dans ce paysage, Schreiber [Schreiber 2000] a suggéré une mesure d'entropie de transfert (TE : Transfer Entropy). Cette notion n'introduit pas de modèle de génération des signaux pour quantifier la part d'information produite dans un système imputable à celle produite antérieurement par un second système. Cette technique utilise des probabilités de transition temporelle entre valeurs d'état, l'espace des états d'un signal étant défini à chaque instant par les valeurs possibles d'un vecteur constitué d'un certain nombre d'échantillons du passé. Cette méthode a été testée sur des modèles non linéaires de "référence" (Rössler, Lorenz) et sur signaux EEG réels, mais sans référence terrain [Sabesan 2007, Sabesan 2009a, Sabesan 2009b]. De par les propriétés qu'elle partage avec l'information mutuelle tout en prenant en compte les dynamiques de transfert d'information, cette approche est une candidate tout à fait légitime pour distinguer dans un système les éléments "maître" des éléments "esclave" de même que les asymétries dans les interactions des sous-systèmes. Par ailleurs, une autre technique non linéaire [Lopes Da Silva 1989], utilisant un coefficient de corrélation non linéaire, a été appliquée pour identifier la dépendance entre deux signaux pour l'analyse de signaux EEG [Louis Dorr 2007, Wendling 2010].

2.1.3. Relations entre connectivités structurelle, fonctionnelle et effective

Cette section résume les recherches actuelles à un niveau d'information un peu plus élevé : la relation entre connectivités anatomique, fonctionnelle et effective est aujourd'hui un véritable challenge en neurosciences, la question posée étant de savoir comment des réseaux fonctionnels et effectifs émergent de la connectivité structurelle cérébrale. Cette question sous-tend naturellement les notions de ségrégation et d'intégration évoquées au début de ce chapitre [Friston 1994, Tononi 1994] qui, au premier abord, apparaissent opposées. La ségrégation fait référence à l'existence de neurones spécialisés au sein des aires corticales, organisés en des populations neuronales distinctes et regroupés pour former des groupes quelque peu isolés. La ségrégation fonctionnelle nécessite une indépendance mutuelle dans la décharge de ces groupes spécialisés de neurones. Le principe complémentaire, l'intégration, provoque l'activation coordonnée de populations neuronales distribuées, et l'intégration fonctionnelle est basée sur la forte cohérence entre leurs activités communes. Cette dualité entre ségrégation et intégration dans les réseaux cérébraux génère de l'information hautement diversifiée à l'origine de graphes informationnels très complexes [Sporns 2000]. Si la connectivité structurelle s'observe habituellement sur des échelles de temps longues (modifications dues à l'âge, à la progression d'une maladie, ...), les connectivités fonctionnelle et effective s'examinent sur des échelles temporelles beaucoup plus réduites (fluctuations rapides observées par exemple lors de tâches sensorielles ou cognitives). Néanmoins les perturbations enregistrées sur ces durées courtes n'affectent généralement pas la structure topologique globale. La difficulté est donc de pouvoir mettre en correspondance des patterns révélateurs de ces différentes connectivités. Si les patterns de connectivité structurelle sont de réelles contraintes pour les dynamiques des circuits corticaux, capturées par les connectivités de plus haut niveau, les fluctuations temporelles rapides relevées à ce plus haut niveau peuvent révéler des changements dans des variables physiologiques sans nécessairement modifier la connectivité de plus bas niveau (*i.e.* structurelle) et de là provient tout la difficulté de représentation et de mise en correspondance.

Les modèles computationnels offrent une réponse partielle à cette question de mise en relation "structure – fonction" dans les réseaux cérébraux. Des simulations de réseaux à grande échelle ont mis en évidence l'émergence de diagrammes de connectivité spatio-temporels complexes à différentes échelles d'observation [Honey 2007]. Lorsque la connectivité fonctionnelle est estimée sur des fenêtres temporelles longues, la topologie des réseaux structurel et fonctionnel est identique, mais les réseaux fonctionnels estimés sur des horizons plus courts sont moins fortement contraints par le câblage structurel. Si la majorité des études supportent l'idée que les réseaux structurels ont un certain impact sur les réseaux fonctionnels sur de longues périodes d'estimation, il n'est pas aussi évident de savoir comment la topologie structurelle intervient dans la reconfiguration de réseaux fonctionnels et est elle-même remodelée au regard de la plasticité fonctionnelle, ceci sur des échelles temporelles plus réduites. Analyser parallèlement les différentes cartes de connectivité enregistrées dans des conditions variées (au repos ou dans la réalisation de tâches spécifiques) s'avère nécessaire pour mieux comprendre les relations entre ces multiples connectivités.

2.2. Modèles neurocomputationnels de l'épilepsie

Puisqu'une crise épileptique peut se décrire comme un désordre neuronal, comprendre l'émergence de ce désordre et pouvoir prédire les crises sont d'extraordinaires défis que des scientifiques ont à cœur de relever via l'élaboration de modèles computationnels analytiques [Touboul 2011]. Même si les mécanismes exacts conduisant aux différentes formes d'épilepsie sont encore mal connus, il est admis que la nature des interactions entre neurones et les propriétés des neurones eux-mêmes sont altérés dans les réseaux épileptogènes. La plupart des études s'accordent à penser que l'épilepsie est souvent liée à une hyperexcitabilité et une hypersynchronisation des réseaux impliqués. L'investigation de telles altérations est essentielle pour la compréhension de cette pathologie et peut avoir une implication forte dans le développement de futurs traitements. La littérature foisonne de modèles physiologiques computationnels qui s'étendent de l'échelle microscopique (description au niveau cellulaire) à l'échelle macroscopique (activité de réseaux corticaux). Indiquons dès à présent que, dans le cadre de cette thèse, seuls les modèles macroscopiques seront considérés.

2.2.1. Modèles microscopiques

Nous ne donnons ici qu'un aperçu très élémentaire de modèles microscopiques, qui ne sont pas par ailleurs ceux utilisés pour valider les méthodes étudiées dans cette thèse. Malgré les contraintes liées aux puissances de calculs et les incertitudes de détails des connexions neuronales et des propriétés biophysiques interneuronales, les modèles microscopiques restent très utiles dans le champ d'étude de l'épilepsie. Le cadre général de ces modèles consiste à reproduire des données expérimentales, acquises sur des capteurs de champs électriques ou magnétiques ou à partir d'autres modalités comme l'imagerie, par ajustement de paramètres de réseaux neuronaux et intraneuronaux et à examiner l'effet de différents facteurs sur le comportement de ces réseaux. Ces modèles peuvent aller de la représentation d'une simple synapse ou d'un seul neurone à des réseaux composés de millions de neurones. Le caractère rédhibitoire du coût de calculs justifie l'effort porté par les chercheurs sur des modèles faisant intervenir seulement quelques cellules. Néanmoins, les modèles avec peu de cellules restent utiles pour établir des prédictions sur des réseaux plus conséquents quantitativement. Par exemple, Skinner *et al.* [Skinner 2005a, Skinner 2005b] ont proposé un modèle à deux cellules qu'ils ont généralisé à plusieurs dizaines de neurones en conservant le même paramétrage et observé que ce réseau étendu présentait un graphe d'activation analogue à celui du réseau minimaliste. Ils ont ainsi extrait des paramètres pertinents dans la description d'un état épileptique. Dans la plupart des modèles de crises épileptiques, il semble que le passage d'une inhibition dominante à une excitation dominante soit responsable de la transition d'un état précritique à un état critique. Toutefois, certains travaux ont conduit à des résultats parfois surprenants que l'intuition n'aurait pas suggéré montrant l'intérêt des modèles computationnels en épilepsie [Van Drongelen 2005, Van Drongelen 2007]. Généralement, les études sur les modèles microscopiques ont été utilisées pour représenter une activité neuronale locale dans une région précise du cerveau supposée présenter une organisation isotrope. Néanmoins, des modèles microscopiques prenant en considération l'interdépendance entre deux ou plusieurs régions, introduisant ainsi une forme d'anisotropie structurelle, ont également été considérés [Destexhe 1996] et ont permis d'apprécier le rôle de certaines structures dans la génération de certains types d'épilepsie. De même, des modèles

détaillés ont aidé dans la mise en évidence d'oscillations rapides apparaissant juste avant le démarrage de crises spontanées [Traub 1982, Traub 2001, Traub 2003, Traub 2005].

2.2.2. Modèles macroscopiques

Le cerveau est un système complexe non linéaire difficile à modéliser mathématiquement, nécessitant des millions de paramètres et variables d'état. Afin de reproduire l'activité globale qui peut être enregistrée par des électrodes extracellulaires, des chercheurs ont développé des modèles macroscopiques [Ermentrout 1998, Faugeras 2009], justifiés *a priori* par le fait que les neurones sont organisés dans des populations homogènes différentes partageant des caractéristiques communes.

Les premiers travaux remontent aux années 70 avec les premières équations de Wilson-Cowan [Wilson 1972,1973] et les premiers résultats de Mountcastle [Mountcastle 1957], Hubel et Wiesel [Hubel 1963,1965,1968] qui fournirent les premières preuves de l'existence de populations macroscopiques. Puis suivirent les travaux de Freeman sur le système olfactif du chat [Freeman 1973, 1975, 1987], et de Lopes Da Silva sur le rythme alpha chez le chien [Lopes Da Silva 1974, Lopes Da Silva 1976]. Aujourd'hui ces études ont trouvé un écho favorable et suscité les développements d'autres équipes [Suffczynski 2001, Suffczynski 2006, Wendling 2000, Wendling 2002, Zetterberg 1978].

Dans les modèles macroscopiques, les signaux (LFP ou EEG) reflètent l'activité globale émergeant d'interactions microscopiques entre des milliers de neurones. Ces activités couplées sont résumées dans l'interaction de variables macroscopiques caractérisant l'activité moyenne de sous-populations neuronales interconnectées (principalement les cellules pyramidales et les interneurones). L'accent est mis ici sur les propriétés des populations neuronales vues comme un tout et non sur celles des cellules individuelles [Touboul 2011]. Une variable importante dans ces modèles est le taux de décharge d'une population (calcul d'un taux moyen sur tous les neurones d'une population donnée) qui résume l'activité de décharge de l'ensemble des neurones dans une sous-population d'un modèle microscopique. Cette variable satisfait souvent une équation différentielle ordinaire non linéaire, stochastique.

Le modèle macroscopique le plus simple est celui à deux sous-populations (excitatrice et inhibitrice) avec des connexions excitatrice et inhibitrice entre elles [Jansen 1993, Jansen 1995], mais des modèles plus sophistiqués avec plus de deux sous-populations s'avèrent nécessaires pour produire des activités EEG plus complexes [Lopes Da Silva 1974, Wendling 2000]. Le modèle retenu dans cette thèse est celui de Wendling *et al* [Wendling 2002]. Dans ce modèle, une troisième sous-population est introduite si l'on considère que différents types de projections inhibitrices vers les cellules pyramidales peuvent intervenir (dans le cas d'épilepsie MLT (Mesial Temporal Lobe) par exemple). Une sous-population d'interneurones projette vers les synapses dendritiques des cellules principales alors que la seconde sous-population d'interneurones projette vers des synapses somatiques. Le modèle montre que la transition d'un état normal à une activité critique se produit quand les deux types d'inhibition (lente et rapide) sont diminués suivant certains rapports. Pour une excitation relativement élevée, le système passe d'une activité de fond à des rythmes rapides quand l'inhibition dendritique lente est réduite tout en conservant une inhibition somatique rapide fixe. Des

décharges rapides de faible amplitude sont observées quand l'inhibition lente est davantage réduite. Finalement, une activité paroxystique de forte amplitude apparaît lorsque l'inhibition dendritique lente est légèrement accrue et l'inhibition somatique rapide décroît.

Un modèle à quatre sous-populations a également fait l'objet d'études [Suffczynski 2004] non détaillées ici et une revue de ces modèles de populations est donnée dans [Lopes Da Silva 2003]. Récemment, une étude de Kramer *et al.* [Kramer 2005, Kramer 2006] s'est inspirée de ce type d'approche pour développer des stratégies de contrôle de l'activité cérébrale humaine dans le but d'empêcher l'apparition de crises.

Les modèles en champ moyen présentent des avantages sur les modèles microscopiques. L'utilité de ces modèles agrégés est souvent présentée comme licite du fait que les macroélectrodes utilisées pour les enregistrements EEG permettent le recueil de potentiels de champs locaux moyens provenant de populations neuronales et sont donc adaptés pour explorer l'activité EEG et examiner les transitions intercritiques/critiques. Il faut noter cependant que ces modèles doivent suffisamment capturer les propriétés de modèles microscopiques plus détaillés (en les agrégeant, bien entendu) afin de rendre compte de la véritable activité cérébrale. Bien souvent, l'activité épileptique couvre des régions étendues et implique plusieurs structures corticales et sous-corticales. Dans un tel contexte, ces modèles sont faciles à analyser numériquement puisqu'ils ne font intervenir qu'un nombre limité de paramètres et de variables. Ils représentent la meilleure alternative pour décrire physiologiquement les processus épileptiques apparaissant dans des systèmes à grande échelle. L'inconvénient majeur de ces modèles est clairement qu'ils ne peuvent rendre compte explicitement des mécanismes cellulaires et/ou moléculaires de l'épileptogenèse, rendant impossible à ce niveau une quelconque modélisation "thérapeutique" (en termes de développement de molécules provoquant des crises), là où les modèles plus microscopiques sont plus adaptés pour la compréhension de tels mécanismes et se positionnent donc avantageusement pour une aide à la pharmacothérapie (pour cibler des modifications moléculaires). Toutefois, bien que la modélisation biophysique joue un rôle fondamental, l'utilisation de tels modèles est limitée par des contraintes de puissances de calculs, d'incertitudes dans la connaissance détaillée des systèmes neuronaux et de simplification requise pour l'analyse numérique. Une approche intermédiaire transversale ("across-scale") établissant des relations entre les variables sous-cellulaires/cellulaires des modèles microscopiques et les paramètres agrégés des modèles macroscopiques s'impose sans doute comme une stratégie prometteuse pour combler les "boîtes noires" qui existent entre ces deux "mondes".

Chapitre 3

Protocole expérimental

Nous présenterons dans le chapitre suivant différentes manières de calculer un indice de connectivité. Etant donné le nombre important d'indices présents dans la littérature et ceux que nous proposons, l'évaluation comparée des performances s'avère d'ores et déjà une tâche difficile. De manière standard, nous avons utilisé des modèles de simulation de signaux et testé également nos algorithmes sur un petit nombre de signaux réels enregistrés sur un cochon d'inde préparé suivant un modèle animal d'épilepsie. Comme largement introduit dans le chapitre 2, l'utilisation pratique de ces indices de connectivité ambitionne de répondre, pour une paire de signaux X_i et X_j , à différentes questions :

- existe-t-il un lien statistique entre X_i et X_j , la négative correspondant à une hypothèse H_0 ?
- si oui, peut-on dire que X_i influence X_j (hypothèse H_{ij}) ou que X_j influence X_i (hypothèse H_{ji}) ou que l'influence est réciproque (hypothèse H_{i-j}) ?

Le premier de ces points correspond à l'existence ou non d'une connectivité fonctionnelle. Le deuxième correspond à la caractérisation en termes de connectivité effective de cette connectivité fonctionnelle, connectivité dont on ignore de toute façon si elle est due ou non à un lien direct (tel qu'on l'entend d'un point de vue anatomique) entre les sous-systèmes générant respectivement les deux signaux.

Ces mêmes questions peuvent s'énoncer conditionnellement à un environnement qui sous-entend la disponibilité d'un troisième signal X_k (ou, dans le principe, d'un ensemble de signaux auxiliaires) :

- conditionnellement à X_k , existe-t-il un lien statistique entre X_i et X_j , la négative correspondant à une hypothèse $H_{0/k}$?

- si oui, et conditionnellement à X_k , X_i influence-t-il X_j (hypothèse $H_{ij|k}$), X_j influence-t-il X_i (hypothèse $H_{ji|k}$) ou a-t-on une influence réciproque (hypothèse $H_{i-j|k}$) ?

Le premier point renvoie à la question de l'existence d'une connectivité fonctionnelle directe et le deuxième à celle d'une connectivité effective directe. Notons que le terme "direct" ne s'entend ici que relativement au contexte défini par X_k (qui, rappelons-le, peut s'étendre dans le principe à un ensemble comportant plus d'un signal scalaire).

Les hypothèses H_0 et $H_{0/k}$ sont très génériques (*i.e.* ne font appel respectivement qu'aux notions d'indépendance simple et conditionnelle). Les autres sont plus subtiles, et requièrent une définition de la notion d'influence, qui peut être la causalité au sens de Wiener et Granger, présentée dans le chapitre 4, et qui nécessite de plus en pratique d'identifier un modèle générique, ou du moins à préciser des ordres de processus supposés Markoviens. Notons que si l'influence effective, dans une direction ou l'autre, n'est recherchée qu'au sens de Granger, elle n'adresse pas la question de connectivité effective instantanée ("instantaneous causality").

Les modèles proposés dans la suite pour les simulations ont donc été choisis de sorte que les réponses (vérité terrain) aux questions posées plus haut paraissent *a priori* non ambiguës. C'est évidemment le cas pour des modèles autorégressifs (AR) linéaires ou non linéaires, et également pour les modèles de populations neuronales (appelés dans ce qui suit modèles physiologiques), puisque leurs équations encodent explicitement les liens de connectivité effective. Notons que les modèles physiologiques utilisés ici sont de même nature (tant d'un point de vue mathématique que sémantique) que les modèles dynamiques utilisés dans l'approche DCM, mais que le seul usage qui en est proposé ici est la simulation de signaux plus réalistes que ceux générés par des modèles AR (linéaires ou non linéaires). Une procédure DCM "classique" procéderait par estimation des coefficients de couplages (réels positifs) tels que décrits dans le paragraphe 3.1.3. Notons deux problèmes pour cette approche. Le premier est qu'elle ne répond à une question sur la connectivité que par des valeurs de coefficients de structure, et beaucoup moins en général sur les propriétés des trajectoires d'état résultantes et donc sur l'intensité d'un lien dynamique de causalité. Si un paramètre de couplage nul implique une absence de causalité, en cas de valeur positive il reste à évaluer la force du lien dynamique induit. De ce point de vue, les approches basées sur les signaux adressent directement la question posée. Le second problème est l'affectation de valeurs aux paramètres autres que ceux de connectivité, *i.e.* les paramètres intra-population. Tout particulièrement, en épilepsie, les paramètres liés à l'excitation et l'inhibition sont *a priori* dépendants du type d'activité et doivent donc être estimés, ce qui n'est pas sans poser de problème [Frogerais 2008, Wendling 2001, Wendling 2005].

Pour résumer, afin de pouvoir juger la pertinence des outils proposés, il convenait de se donner, pour disposer d'une vérité-terrain, et au-delà des modèles AR classiques, des modèles de génération de signaux physiologiquement plausibles. Tous ces modèles qui représentent des instanciations de graphes de propagation variés, que l'on se donne *a priori*, afin de construire des scénarios menant à des patterns de flux d'information variés, sont présentés dans les sections 3.1.1, 3.1.2, 3.1.3.

Enfin, la réponse à la véracité d'une hypothèse testée H_T , concernant une interprétation d'observations $O = (X_i, X_j)$ ou $O = (X_i, X_j, X_k)$, passe classiquement par l'introduction d'un test d'hypothèse qui prend le plus souvent la forme : H_T est acceptée si une certaine statistique $T(O)$, à valeurs réelles, s'avère appartenir ou non à un domaine d'acceptation (un intervalle de \mathbb{R} le plus souvent). Pour les hypothèses portant sur la connectivité simple ou conditionnelle, questionnées ici, $T(O)$ pourra prendre la forme $Ind_{X_i \rightarrow X_j}$ ou $Ind_{X_i \rightarrow X_j / X_k}$ où Ind désigne génériquement un des indices proposés dans le chapitre 4. Des précisions concernant la stratégie adoptée et les aspects statistiques afférents sont données dans la section 3.2.2.

3.1. Bases de données

3.1.1. Modèles linéaires autorégressifs

Nous considérons le modèle suivant de génération de signaux, basé sur une modélisation linéaire autorégressive :

$$\begin{cases} x_1(t) = 0.95\sqrt{2}x_1(t-1) - 0.9025x_1(t-2) + w_1(t) \\ x_2(t) = -0.5x_1(t-1) + 0.25\sqrt{2}x_2(t-1) - \beta x_3(t-3) + w_2(t) \\ x_3(t) = -\alpha x_1(t-2) - 0.5x_2(t-2) - 0.25\sqrt{2}x_3(t-2) + w_3(t) \end{cases} \quad (3.1)$$

où $w_j(t)$, $j = 1, 2, 3$, sont des bruits blancs indépendants de moyenne nulle et de variance unité. Le paramètre β ($\beta = 0.5$) est introduit pour modéliser les flux bidirectionnels entre les signaux x_2 et x_3 . L'introduction du paramètre α permet de considérer deux patterns d'interactions causales, soit des relations directes ($\alpha = 0.5$) soit des relations indirectes ($\alpha = 0$), ce qui conduit explicitement à quatre modèles comme on le voit sur la Fig. 3.1 (modèle 1 : $\alpha = 0$, $\beta = 0$, modèle 2 : $\alpha = 0.5$, $\beta = 0$, modèle 3 : $\alpha = 0$, $\beta = 0.5$, modèle 4 : $\alpha = 0.5$, $\beta = 0.5$).

3.1.2. Modèles non linéaires autorégressifs

Pour le modèle stochastique non linéaire, les signaux sont gouvernés par les équations suivantes :

$$\begin{cases} x_1(t) = 3.4x_1(t-1)(1 - x_1^2(t-1))e^{-x_1^2(t-1)} + w_1(t) \\ x_2(t) = 3.4x_2(t-1)(1 - x_2^2(t-1))e^{-x_2^2(t-1)} - 0.5x_1(t-1) + 0.25\sqrt{2}x_2(t-1) - \beta x_3(t-3) + w_2(t) \\ x_3(t) = 3.4x_3(t-1)(1 - x_3^2(t-1))e^{-x_3^2(t-1)} - \alpha x_1(t-2) - 0.5x_2(t-2) - 0.25\sqrt{2}x_3(t-2) + w_3(t) \end{cases} \quad (3.2)$$

où $w_j(t)$, $j = 1, 2, 3$, sont des bruits blancs indépendants de moyenne nulle et de variance unité. Comme précédemment, le paramètre β est introduit pour modéliser les flux bidirectionnels entre les signaux x_2 et x_3 . Dans ce cas, ce paramètre est fixé à 0.5. Le paramètre α est introduit pour considérer deux patterns d'interactions causales, ou des relations directes ($\alpha = 0.5$) ou des relations

indirectes ($\alpha = 0$), ce qui conduit, comme précédemment, à quatre modèles (modèle 1 : $\alpha = 0$, $\beta = 0$, modèle 2 : $\alpha = 0.5$, $\beta = 0$, modèle 3 : $\alpha = 0$, $\beta = 0.5$, modèle 4 : $\alpha = 0.5$, $\beta = 0.5$) (cf. Fig. 3.1).

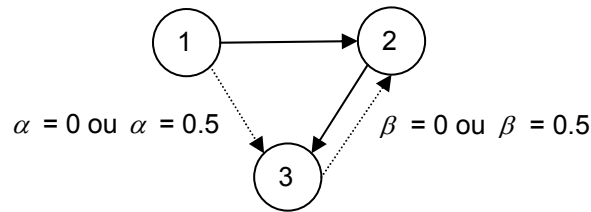


Fig. 3.1 – Modèle de simulation des signaux.

3.1.3. Modèles physiologiques

Dans un second temps, nous avons utilisé un modèle du type EDS (équations différentielles stochastiques) à temps continu simulé en temps discret. Il est basé sur la physiologie et a été introduit dans [Wendling 2005] pour représenter les signaux de potentiel, chacun enregistré à proximité d'une population neuronale (les différentes populations étant espacées mais potentiellement couplées). Chaque modèle de population simule un potentiel membranaire moyen local qui est converti en un signal intracérébral EEG par une fonction de transfert quasi-statique [Wendling 2001, Wendling 2005]. Dans le modèle ainsi retenu, chaque population est constituée de trois sous-populations de neurones qui interagissent mutuellement : une sous-population de neurones principaux excitateurs et deux sous-populations de neurones inhibiteurs. La sous-population principale excitatrice exerce une action en retour sur elle-même. Les cellules de la première sous-population inhibitrice correspondent à des interneurons qui se projettent sur la région dendritique de la population principale. Concernant la seconde, il s'agit d'interneurones inhibiteurs pour lesquels les contacts synaptiques s'effectuent dans la région somatique des neurones principaux. La description mathématique du modèle est donnée dans [Frogerais 2008]. Puisque les cellules pyramidales sont des neurones excitateurs qui projettent leurs axones vers d'autres aires du cerveau, le modèle tient compte de cette organisation en utilisant le taux de décharge moyen des potentiels d'action issus des cellules principales d'une population i comme entrée des cellules principales d'une seconde population j . Cette connexion entre les populations i et j est ajustable par un paramètre K^{ij} proportionnel au nombre de liens synaptiques efficaces pour un type d'activité donné. Une modification appropriée des coefficients K^{ij} permet de construire des systèmes dans lesquels les populations peuvent être couplées de manière uni- ou bidirectionnelle. Les autres paramètres de ce modèle sont des paramètres internes aux populations elles-mêmes. Ils comprennent les gains excitateurs et inhibiteurs dans les boucles de rétroaction ainsi que les coefficients liés au nombre de contacts synaptiques entre les sous-populations. Ils sont ajustés pour contrôler le comportement intrinsèque de chaque population (activité de fond vs activité épileptique). Nous pouvons reproduire les mêmes types de connexions que ceux de la Fig. 3.1 et dériver ainsi 4 scénarios. Dans le scénario pour lequel nous présenterons ultérieurement des résultats, la population 1 entraîne les populations 2 et 3, ces deux populations étant couplées bidirectionnellement. Si nécessaire, en référence aux modèles décrits dans les deux sections

précédentes, ce scénario sera noté "modèle 4". Les valeurs choisies pour les paramètres sont $K^{12} = K^{13} = K^{32} = K^{23} = 1500$ permettant de rendre les trois populations épileptiques.

3.1.4. Signaux réels

Les activités épileptiformes sont induites par la perfusion de bicuculline dans le cerveau isolé de cochon d'inde. Des enregistrements simultanés sont opérés dans le cortex piriforme, le cortex entorhinal médian et latéral, l'aire CA1 de l'hippocampe, et le cortex périrhinal. Pour tester nos approches, nous disposons d'enregistrements sur une fenêtre temporelle d'une dizaine de secondes correspondant à une phase d'activité rapide (FOA : Fast Onset Activity) apparaissant lors de l'installation d'une crise.

3.2. Méthodologie d'évaluation

3.2.1. Détermination de graphes

Dans un premier temps, il s'agit de calculer des indices de connectivité effective sur les ensembles de signaux définis dans la section précédente puis, à partir des valeurs obtenues, d'en déduire des graphes de propagation d'information. Au-delà de la simple observation visuelle, la difficulté réside dans la définition d'un seuil objectif au-dessus duquel on décidera qu'il y a connectivité effective, autrement dit de l'introduction d'un test pour chaque paire (X_i, X_j) (analyse non conditionnée) ou chaque triplet (X_i, X_j, X_k) (analyse conditionnée). Tous les indices $Ind_{X_i \rightarrow X_j}$ et $Ind_{X_i \rightarrow X_j / X_k}$ étant théoriquement nuls sous l'hypothèse H_0 et à valeur croissante quand le lien (fonctionnel ou effectif) se renforce, la forme du test sera trivialement, respectivement dans les cas non conditionnel et conditionnel :

- $T_2(X_i, X_j) = Ind_{X_i \rightarrow X_j} > \lambda \Rightarrow H_0$ rejetée et H_{ij} acceptée
- $T_3(X_i, X_j, X_k) = Ind_{X_i \rightarrow X_j / X_k} > \lambda \Rightarrow H_0$ rejetée et $H_{ij/k}$ acceptée

La forme de T_2 ou T_3 étant donnée, il reste à choisir le seuil λ . L'approche communément admise est de choisir λ tel que $P(T_2 > \lambda / H_0) \leq pfa$ ou $P(T_3 > \lambda / H_{0/k}) \leq pfa$, où pfa est une probabilité d'erreur de première espèce (dite aussi probabilité de fausse alarme, ou probabilité de faux négatif) fixée par l'utilisateur. Toutefois, le calcul théorique de ces quantités n'est accessible que dans des cas simples, où pour de plus ou moins bonnes raisons asymptotiques la loi de T_2 ou T_3 sous l'hypothèse nulle devient invariante relativement à la statistique des signaux analysés. En pratique, il est généralement nécessaire d'utiliser une distribution empirique de T_2 ou T_3 obtenue par une technique de type génération de données de substitution ("surrogate data"), explicitée pour ce qui nous concerne dans la section qui suit.

Pour un triplet de signaux donné (i, j, k) , un graphe de propagation est obtenu de la façon suivante :

pour $(i, j) = (1, 2), (2, 1), (1, 3), (3, 1), (1, 2), (2, 1)$:

1) déterminer $\lambda_{i,j}$ (resp. $\lambda_{i,j/k}$)

2) calculer $Ind_{X_i \rightarrow X_j}$ (resp. $Ind_{X_i \rightarrow X_j/X_k}$)

3) si $Ind_{X_i \rightarrow X_j} > \lambda_{ij}$, accepter H_{ij} (resp. si $Ind_{X_i \rightarrow X_j/X_k} > \lambda_{ij}$, accepter $H_{ij/k}$) (et sinon

rejeter)

fin

L'étape (1) correspond au calcul d'un seuil adaptatif fonction des signaux observés. Néanmoins, il est envisageable d'obtenir ce seuil en amont de la boucle au cours d'une étape préalable utilisant une base d'apprentissage.

3.2.2. Analyse des distributions sous H_0

La situation face à laquelle nous sommes confrontés est complexe. En effet, pour un jeu de trois populations correspondant aux signaux x_1 , x_2 et x_3 (réalisation des signaux aléatoires X_i , $i = 1, 2, 3$) pour lesquels nous mesurons un indice de direction $Ind_{X_i \rightarrow X_j}$ d'une population i vers une population j , étant entendu que nous disposons d'un modèle de connectivité reliant ces trois populations, on aura compris que le problème est d'évaluer l'écart par rapport à une hypothèse H_0 dans laquelle les signaux X_i et X_j seraient indépendants la difficulté étant d'obtenir une distribution théorique de $Ind_{X_i \rightarrow X_j}$ sous H_0 ou de $Ind_{X_i \rightarrow X_j/X_k}$ sous $H_{0/k}$. Cette difficulté peut être levée en recourant à des "données de substitution" ("surrogate data") synthétisées à partir des données originales, et qui garantissent leur indépendance pour ainsi disposer d'une statistique de référence sous H_0 (ou $H_{0/k}$). Il faut pour cela définir une stratégie pour modifier les réalisations x_i ou x_j de telle sorte qu'elles correspondent à deux signaux indépendants de mêmes caractéristiques fréquentielles marginales (qui, on le sait, influencent largement la variance de toute statistique fonction des observations). Nous avons envisagé ici deux méthodes décrites ci-après.

1^{ère} méthode

Considérant deux réalisations x_i et x_j issues de signaux X_i et X_j initialement dépendants, on calcule la transformée de Fourier de l'une d'elles, par exemple x_j , on passe au signal analytique, puis on remplace la phase de ce signal, pour chaque fréquence discrète, par une valeur tirée aléatoirement sur $[0, 2\pi]$ et, après transformée de Fourier inverse, on garde la partie réelle du signal obtenu. On construit ainsi un processus X'_j interprétable comme étant indépendant de X_i , ou du

moins décorrélé, et qui a la même densité spectrale de puissance que X_j . Cette technique des phases aléatoires permet ainsi de générer des signaux décorrélés dont les caractéristiques spectrales (caractérisation au second ordre) marginales sont conservées. Pour obtenir une distribution statistique de $Ind_{X_i \rightarrow X_j}$ sous H_0 , ou de $Ind_{X_i \rightarrow X_j/X_k}$ sous $H_{0/k}$, il suffit alors de répéter l'opération pour un nombre suffisant de réalisations indépendantes x'_j (de X'_j). Un seuil peut être alors obtenu par calcul du quantile correspondant à une probabilité pfa désirée. Cette stratégie peut être utilisée aussi bien pour des signaux réels que pour des signaux simulés et permet de calculer le seuil pour une paire donnée de signaux (calcul du seuil à l'intérieur de la boucle dans l'algorithme donné précédemment).

2^{ème} méthode

Dans la première méthode, la destruction de la phase "modifie" quelque peu les caractéristiques temporelles "non linéaires" de ces signaux, ce qui peut être regrettable. Une alternative pour construire des signaux sous l'hypothèse H_0 (ou $H_{0/k}$) en évitant cet inconvénient est décrite ci-après.

Considérons une suite de M réalisations indépendantes (x_i^m, x_j^m, x_k^m) , $m = 1, \dots, M$, obtenues avec un même modèle (même structure et mêmes paramètres) ou par répétition d'une expérience contrôlée, par exemple sur modèle animal. Intéressons-nous à la relation entre les signaux X_i et X_j . Pour fabriquer des paires (X_i, X'_j) indépendantes, avec conservation intégrale des lois marginales, il suffit de choisir (au hasard ou non) des paires $(X_i^m, X_j^{m'})$, $m \neq m'$, en nombre suffisant pour atteindre la précision statistique voulue. Pour cette deuxième méthode, la détermination d'un seuil peut se faire en amont de la boucle dans l'algorithme présenté plus haut.

Ainsi, pour valider nos approches sur signaux simulés (modèles AR linéaire et non linéaire et modèle physiologique), nous calculerons nos indices de connectivité sous les hypothèses H_0 (ou $H_{0/k}$) suivant l'une des méthodologies précédentes. Cette étude sera cependant restreinte aux indices *a priori* les plus performants (donnant des graphes de propagation cohérents) et ce dans le cas d'une connectivité particulière pour l'ensemble de ces signaux. Pour chaque indice, il s'agira de comparer les résultats obtenus sous H_0 (ou $H_{0/k}$) avec ceux obtenus avec les signaux observés (non modifiés).

En ce qui concerne les signaux réels, n'ayant manipulé pour cette étude qu'un petit nombre d'entre eux, nous retiendrons l'approche par données de substitution en rendant aléatoire la phase des signaux.

Chapitre 4

Méthodes

Comme nous l'avons dit précédemment, nos travaux ont pour objectif final d'analyser les signaux électroencéphalographiques intracérébraux enregistrés lors de crises d'épilepsie et plus particulièrement d'identifier les structures cérébrales impliquées dans les différentes phases d'une crise, quantifier l'information portée par les différentes observations et investir le plus finement possible les relations inter-structures qui s'établissent au cours des épisodes épileptiques.

Investir la notion de connectivité cérébrale sous-tend souvent l'appréhension successive des étapes de spécification de modèle, d'identification et d'inférence causale [Valdes-Sosa 2011]. A ce niveau, il convient de préciser que nous ne considérons pas de modèles spécifiques utilisés pour mesurer des relations causales entre signaux bien que notre propos ne soit pas tant basé sur la connectivité fonctionnelle mais sur la connectivité effective. En d'autres termes, nous focalisons nos approches sur la mise en évidence d'une connectivité effective qui relève de la détection générique de relations causales entre systèmes neuronaux. Insistons dès à présent sur le fait que la causalité est un concept épistémologique qui est particulièrement difficile à appréhender via des équations. A cet égard, la notion de causalité peut être pensée en ces termes :

- elle peut être vue comme une précédence temporelle, *i.e.* les causes précèdent leurs conséquences ;
- elle peut être envisagée sous l'angle d'une influence physique, *i.e.* modifier les causes modifie les conséquences.

Cette distinction est importante puisqu'elle est à la base de n'importe quelle détection statistique d'une influence causale. Dans le contexte de la connectivité cérébrale, identifier des relations causales entre deux régions du cerveau se mesure alors soit en termes d'amélioration de la capacité de prédiction d'événements neuronaux distincts temporellement soit en termes d'évaluation de l'effet à distance de modifications sur les événements.

Dans ce chapitre, nous commençons par rappeler le concept de causalité au sens de Wiener-Granger et présentons différentes variantes de l'indice correspondant aussi bien dans le domaine

temporel que fréquentiel. Puis, nous proposons de nouveaux indices basés sur une mesure récente de pente de phase, faisant intervenir différentes fonctions de cohérence afin de répondre aux questions de relations directes vs indirectes et d'éventuelles bidirectionnalités. Nous poursuivons ce chapitre par l'introduction d'une mesure d'entropie de transfert et proposons une nouvelle stratégie pour identifier l'ordre des modèles sous-jacents. Finalement, ce concept est étendu au cas trivarié (étant entendu qu'il pourrait l'être au cas multivarié mais celui-ci ne sera pas envisagé dans le cadre de ce travail).

4.1. Indice de causalité de Granger

Comme précisé dans le chapitre 2, cette approche stipule qu'il existe un lien de causalité d'un signal X_1 vers un signal X_2 si, relativement à l'instant t , la connaissance du passé de X_1 et X_2 améliore la variance d'erreur de prédiction de $X_2(t)$ par rapport à la connaissance du passé de X_2 uniquement [Granger 1969, Wiener 1956]. En 1982, Geweke proposa une série de mesures de causalité basées sur la décomposition spectrale de processus ARX multivariés permettant d'exprimer la causalité de Granger temporelle dans le domaine fréquentiel [Geweke 1982]. Dans le domaine temporel, on s'intéresse ici plus particulièrement aux approches WGCI-P (WGCI-P : Wiener Granger Causality Index in Pairwise analysis, pour des comparaisons par paires), et WGCI-C (WGCI-C : Conditional Wiener Granger Causality Index, pour des comparaisons dans le cas multivarié). De même, dans le domaine fréquentiel, nous nous focalisons sur leurs "homologues" à savoir les indices FGCI-P (pour des comparaisons par paires) et FGCI-C (pour des comparaisons dans le cas multivarié).

4.2. Indices basés sur la pente de phase

Récemment, un indice de pente de phase, noté PSI, basé sur la pente de phase de la fonction de cohérence ordinaire calculée entre deux signaux a été proposé par Nolte [Nolte 2008] pour détecter le flux d'informations dans des graphes de propagation unidirectionnels. L'hypothèse de base de cette approche repose sur l'exploitation de la monotonie de phase entre deux signaux qui apparaît si les composantes fréquentielles d'un premier signal précèdent temporellement celles d'un second. Cet indice a été introduit pour résumer l'information de pente de phase de l'interspectre calculé entre ces deux signaux. L'idée sous-jacente est une représentation du retard entre les signaux observés sur les bandes de fréquences où la cohérence entre ces signaux est significative. L'amplitude de la cohérence permet de pondérer la différence de phase entre deux fréquences consécutives et ainsi de minimiser son impact lorsque l'un des deux facteurs (ou les deux) est (sont) faible(s). Le signe du PSI indique la direction du flux d'information et son amplitude s'accroît avec le retard et la valeur de la cohérence. Cet indice PSI, tel que défini initialement, présente deux inconvénients majeurs : (i) il ne peut se justifier que dans des situations de flux unidirectionnels de par son caractère symétrique, (ii) il échoue dans la discrimination de relations directes et indirectes. En effet, si un troisième signal est à l'origine de la relation linéaire entre deux autres signaux, l'amplitude de la cohérence entre ces deux signaux est unité ce qui ne permet pas de distinguer liaisons directe et indirecte entre ces signaux.

Pour corriger cet effet, nous proposons de remplacer la cohérence ordinaire par la cohérence partielle [Yang 2010], qui, partant de Q signaux, traduit le niveau de couplage entre deux de ces signaux lorsque l'influence des $Q - 2$ autres signaux est prise en compte. Nous définissons ainsi un nouvel indice de causalité noté CI-PC (CI-PC : Causality Index – Partial Coherence). Au-delà de cette approche, pour répondre aux questions de bidirectionnalité, nous introduisons de nouveaux indices de causalité afin de détecter et de différencier relations uni- et bi-directionnelles : un premier est basé sur la cohérence dirigée (DCOH : Directed COHerence) [Saito 1981] et considère une analyse par paires (CI-DCOH), un second sur la fonction de transfert dirigée (DTF : Directed Transfer Function) [Kamiński 1991] suivant une analyse multivariée (CI-DTF) [Yang 2011]. Alors que les cohérences ordinaire et partielle focalisent sur l'information mutuelle entre structures, la cohérence dirigée et la fonction de transfert dirigée se réfèrent davantage au concept de causalité de Granger. Par conséquent, contrairement aux deux cohérences précédentes, ces deux fonctions sont des quantités asymétriques. Introduit par Saito et Harashima [Saito 1981], le concept de cohérence dirigée permet d'analyser conjointement l'information présente dans deux observations, chacune disposant de sa propre source de bruit blanc, considérée comme un processus d'innovation local, et d'une source commune, vue comme un processus d'innovation externe. Ce processus de génération de signaux peut être représenté par un modèle autorégressif bivarié. Alors que la cohérence mesure le degré de corrélation et le considère comme un tout, les "cohérences dirigées" peuvent être assimilées à des "corrélations avec direction" entre les deux signaux observés exprimées dans le domaine fréquentiel et, de ce fait, être perçues comme deux facteurs pondérés contribuant à l'expression de la cohérence globale. Ainsi, étant donné deux signaux observés, représentant les sorties du système étudié, ces cohérences décrivent la connexion entre l'innovation du premier processus (resp. du second processus) et la sortie du système correspondant à la seconde observation (resp. la première observation). Pour formuler ce problème "d'influences dirigées corrélées", Saito et Harashima [Saito 1981] considèrent un processus autorégressif bivarié incluant une source commune de bruit. De là, nous dérivons un indice de causalité asymétrique permettant de détecter des flux bidirectionnels qui n'est plus directement lié à la pente de phase entre les observations elles-mêmes mais à celle estimée entre les sources de bruit et les signaux observés. Suivant le concept de cohérence dirigée, la fonction de transfert dirigée a été introduite par Kamiński et Blinowska [Kamiński 1991] pour traiter un nombre d'observations supérieur à deux. Dans ce cas, contrairement à la précédente approche, il n'y a pas de source commune de bruit et chaque observation est produite par son propre processus d'innovation linéairement combiné avec des versions retardées des autres observations. Comme précédemment, la fonction de transfert entre une entrée et une sortie du système peut être évaluée. Ainsi, dans le cas multivarié, nous pouvons étendre le concept d'indice de causalité à la fonction de transfert dirigée. Suivant les mêmes développements, Baccalá et Sameshima [Baccalá 2001] ont "opposé" la cohérence dirigée partielle (PDC : Partial Directed Coherence) à la cohérence dirigée afin de fournir une information structurelle directe dans le cas multivarié, de la même façon que l'on peut "opposer" la cohérence partielle à la cohérence ordinaire dans le cas de modèles à flux unidirectionnels. Aussi, pour supporter les situations bidirectionnelles pouvant présenter également

des relations directes ou indirectes, introduisons-nous la cohérence partielle dirigée dans un nouvel indice, noté CI-PDC [Yang 2012].

4.3. Entropie de transfert et entropie de transfert conditionnée

Une approche non paramétrique du problème consiste en celle de la théorie de l'information dans laquelle on peut introduire différents indicateurs, comme l'information dirigée. Une autre caractérisation de nature entropique est l'entropie de transfert (TE : Transfer Entropy) proposée par Schreiber [Schreiber 2000]. Celle-ci n'introduit pas de modèle de génération des signaux pour quantifier la part de l'information produite dans un système imputable à celle produite par un second système dans le passé. Cette technique utilise des probabilités de transition temporelle entre valeurs d'état. Une valeur d'état est introduite en chaque instant pour chaque signal. L'espace des états, pour un signal, est défini en chaque instant par un vecteur constitué d'un certain nombre d'échantillons récents. Dans cette technique, certains paramètres de calibration jouent un rôle crucial dans la détermination de la direction du flux d'information entre les deux systèmes mis en jeu – le système maître et le système esclave –, comme les ordres des processus markoviens associés aux systèmes. Dans [Schreiber 2000], seul l'ordre du processus de Markov du système maître est évoqué et choisi identique à celui du système esclave ou imposé à un pour des raisons de calculs. Dans [Sabesan 2007, Sabesan 2009b], ce point-clé est également examiné et deux mesures, l'information mutuelle et la fonction d'autocorrélation [Fraser 1986, Martinerie 1992], sont suggérées pour l'estimation de l'ordre du processus de Markov associé au système esclave, l'ordre du processus maître étant fixé à un, étant présumé que l'état courant du système maître est suffisant pour susciter des changements dans la dynamique du second système. Une alternative est d'utiliser des critères standards, tels que les critères AIC (Akaike Information Criterion) [Akaike 1973] ou BIC (Bayesian Information Criterion) [Schwarz 1978] pour estimer l'ordre des processus. Toutefois, ces critères ne rendent qu'un ordre commun. Or, dans la plupart des situations, les ordres des processus associés aux systèmes diffèrent. Par conséquent, notre stratégie pour améliorer l'estimation de la mesure d'entropie de transfert consiste à généraliser ces deux critères en préconisant un algorithme "glouton", le temps de calcul des différents ordres pouvant devenir prohibitif, pour tendre vers une solution optimale (introduction des critères gAIC et gBIC). D'autre part, cette technique entropique caractérise les interactions causales suivant une analyse des signaux par paires, et ne peut détecter les flux directs et indirects dans des systèmes complexes multivariés. Aussi considérons-nous une extension de cette approche, dénommée CTE (Conditional Transfer Entropy) afin de discriminer les différents types de flux et prendre en compte l'éventuelle contribution d'un troisième signal sur les deux signaux analysés. La même démarche que celle développée dans le cas de la mesure TE est adoptée quant à la détermination des ordres et à la mise en œuvre d'un algorithme "glouton" pour réduire la complexité et tendre vers une résolution optimale. Soulignons que cette approche peut se généraliser, sans difficulté, à la prise en compte de multiples observations.

4.4. Discussion

Dans ce chapitre, nous avons successivement développé des approches issues de la causalité au sens de Granger avant de nous focaliser sur la proposition d'indicateurs dérivés d'une technique trouvant son origine dans l'estimation de la pente de phase entre signaux observés puis de nous orienter vers des approches plus complexes mais sans doute plus prometteuses dans le paysage de signaux présentant des non linéarités. Ces techniques sont testées dans le chapitre suivant, tant sur des modèles simulés simples (linéaires et non linéaires) que sur des modèles physiologiques réalistes mais aussi sur une base réduite de données réelles.

Chapitre 5

Résultats expérimentaux

Les techniques d'analyse de connectivité cérébrale décrites dans le chapitre 4 ont d'abord été testées sur données simulées générées par des modèles vectoriels AR linéaires. A l'issue de cette étape, certaines mesures ont été sélectionnées pour être appliquées sur des modèles AR non linéaires ainsi que sur des données issues d'un modèle physiologique et, pour finir, sur des signaux réels. En simulation, chaque couple (modèle, méthode) étudié a été évalué au moyen de deux cents réalisations de Monte Carlo indépendantes, chaque réalisation correspondant à 3 signaux de 2048 points. Le Tableau 5.1 résume l'ensemble des situations pour lesquelles des résultats sont rapportés dans la version anglaise de ce document. Les résultats d'expérimentation ne sont pas tous rapportés avec la même finesse d'analyse : les modèles AR linéaires se sont vus appliquer le plus grand nombre de méthodes, seul un sous-ensemble d'entre elles ont été retenues pour les modèles AR non linéaires et enfin, sur modèle physiologique et signaux réels, ne sont donnés de résultats que pour les méthodes WGCI-C, CI-PDC et CTE. Dans ce tableau, la lettre T correspond à une présentation sous forme d'un tableau à 6 cases. Chaque case se réfère à l'un des 6 liens analysés et contient une paire (moyenne, écart-type) résumant la distribution statistique de l'indice de connectivité considéré pour ce lien. La lettre F correspond aux résultats sous forme de figures et relatifs aux méthodes fréquentielles, qui passent toutes par le calcul d'un indice fonction de la fréquence. Dans ce cas, des "tableaux" à 6 cases sont également utilisés, chaque case contenant le tracé de la valeur moyenne (toujours sur les 200 réalisations) de cet indice en fonction de la fréquence. Enfin, BP indique qu'une analyse sous forme de boîtes à moustaches ("box plot") a été réalisée, complétée par des courbes COR (courbes opérationnelles de réception reliant la probabilité expérimentale de détection vraie à la probabilité expérimentale de fausse alarme) dans le cas du modèle physiologique. Les principaux résultats et les commentaires associés sont reproduits dans la présente version. Pour les signaux réels, seules des boîtes à moustaches correspondant à l'hypothèse H_0 ont pu être obtenues, puisque l'on ne disposait pour chaque structure enregistrée que d'une réalisation (sur un intervalle temporel positionné au niveau de l'activité rapide) de signal expérimental.

Méthodes	Modèles linéaires				Modèles non linéaires				Modèles Physiol.				Sig. réels
	M1	M2	M3	M4	M1	M2	M3	M4	M1	M2	M3	M4	
WGCi-P	T	T	T	T									
WGCi-C	T	T	T	T/BP	T	T	T	T/BP				BP	BP
FGCi-P	F	F	F	F									
FGCi-C	F	F	F	F	F	F	F	F					
PSI-OC	T	T	T										
CI-PC	T	T											
CI-DCOH	T		T	T			T	T					
CI-DTF	T		T	T			T	T					
CI-PDC			T	T/BP			T	T/BP				BP	BP
TE	T	T	T	T	T	T	T	T					
CTE	T	T	T	T/BP	T	T	T	T/BP				BP	BP

Tableau 5.1. Bases de données et méthodes testées dont les analyses sont données dans la version anglaise du chapitre 5 (Mi se réfère au Modèle i)

T : résultats présentés sous forme de tableaux donnant les moyennes et écarts-types

F : résultats présentés sous forme de figures donnant des valeurs moyennes sur l'échelle des fréquences

BP : résultats présentés sous forme de boîtes à moustaches (Box Plot).

5.1. Analyse des résultats sur modèles AR linéaires

Nous présentons ici une synthèse des résultats, leur analyse détaillée étant consultable dans la section 5.1 du chapitre 5 de la version anglaise.

La méthode WGCi-C atteste d'un bon comportement et permet de retrouver effectivement les quatre modèles dérivés de l'équation (3.1) à la différence de la méthode WGCi-P, qui, travaillant par paires, ne distingue pas nécessairement les relations directes et indirectes. En ce qui concerne les résultats sur les indices fréquentiels, les mêmes conclusions émergent même si l'analyse des résultats est plus complexe dans la mesure où les indices rendent des valeurs par fréquence. Dans tous les cas, si l'on se fie à un contraste visuel souvent évident, ces indices se révèlent capables d'identifier et de discriminer des relations causales unidirectionnelles et bidirectionnelles.

Pour ce qui est des approches basées sur la pente de phase, un premier résultat est en faveur de l'estimation des spectres par modélisation AR face à une estimation par transformée de Fourier, ce qui est naturel étant donné le type de modèle utilisé (AR linéaire). Parmi les indices proposés, l'indice basé sur la cohérence dirigée partielle (CI-PDC) est le plus performant puisque, non seulement il détecte les relations unidirectionnelles et bidirectionnelles, mais il est également capable de "différencier" les connexions directes et indirectes.

Concernant les approches par entropie de transfert, l'étape préliminaire d'optimisation des ordres des modèles a montré la pertinence du critère gBIC pour une estimation plus précise et plus robuste de l'indice de transfert d'entropie. D'autre part, comparativement à TE, l'approche CTE permet de lever l'ambiguïté des relations directes vs indirectes. Les graphes de propagation estimés pour les quatre modèles retenus sont tout à fait cohérents avec les graphes simulés.

En termes d'analyse statistique, sont représentés sur la Fig. 5.1 les résultats obtenus par les indices WGCi-C, CI-PDC et CTE pour le modèle linéaire correspondant aux valeurs suivantes de paramètres dans l'équation (3.1) : $\alpha = 0.5$ et $\beta = 0.5$ (modèle 4). Sans aucune équivoque, quelle que soit l'approche testée, les résultats permettent de distinguer les hypothèses H_0 et \bar{H}_0 pour les liaisons $1 \rightarrow 2$, $1 \rightarrow 3$ et $2 \rightleftharpoons 3$, et donc d'attester de ces connectivités effectives qui sont celles imposées par le modèle. De plus, pour l'indice CI-PDC, il "quantifie" d'une certaine manière les retards imposés. En effet, si l'on compare les liaisons $3 \rightarrow 2$ et $2 \rightarrow 3$, on note un rapport de 3/2 entre les deux qui correspond au rapport des retards respectifs du signal 3 sur le signal 2 et du signal 2 sur le signal 3.

5.2. Analyse des résultats sur modèles AR non linéaires

Comme précédemment, nous ne donnons ici qu'une analyse succincte des résultats, l'ensemble des analyses se trouvant dans la section 5.2 du chapitre 5 de la version anglaise.

Pour le modèle stochastique non linéaire retenu gouverné par l'équation (3.2), les indices de Granger conditionnés, que ce soit en temporel (WGCi-C) ou en fréquentiel (FGCi-C), s'avèrent tout à fait pertinents.

De même, pour les mesures élaborées à partir de l'indice de pente de phase, on peut à nouveau reconnaître la robustesse et la supériorité de l'indice basé sur la cohérence dirigée partielle (CI-PDC) qui rend compte des graphes exacts de propagation quelle que soit la configuration testée. Les relations directes/indirectes et les relations bilatérales sont parfaitement identifiées.

Pour finir, les approches sur l'entropie donnent également des résultats comparables à ceux obtenus dans le cas linéaire, à savoir qu'elles sont à même de détecter des influences bidirectionnelles mais, là aussi, seule la mesure CTE relève le défi de discriminer relations directes et indirectes.

A l'instar de la section précédente, nous donnons sur la Fig. 5.2 les résultats d'analyse statistique des méthodes WGCi-C, CI-PDC et CTE pour le même type de connectivité (modèle 4, *i.e.* $\alpha = 0.5$ et $\beta = 0.5$ dans l'équation (3.2)). Comme dans le cas linéaire, quelle que soit l'approche testée, les résultats obtenus permettent de dissiper toute potentielle ambiguïté sur les relations d'influence entre les signaux testés. Les liaisons $1 \rightarrow 2$, $1 \rightarrow 3$ et $2 \rightleftharpoons 3$ sont manifestes et permettent d'authentifier le modèle sous-jacent.

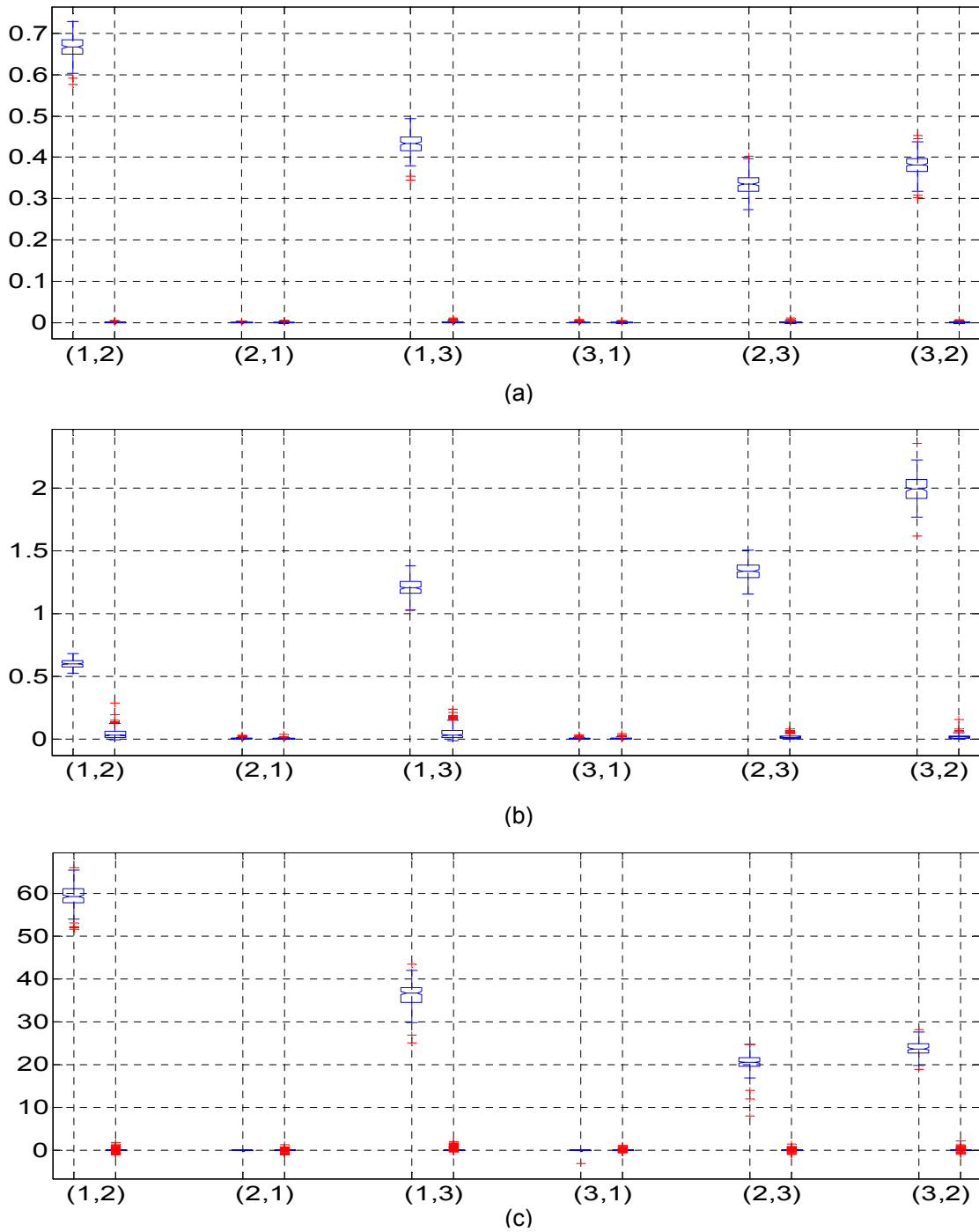
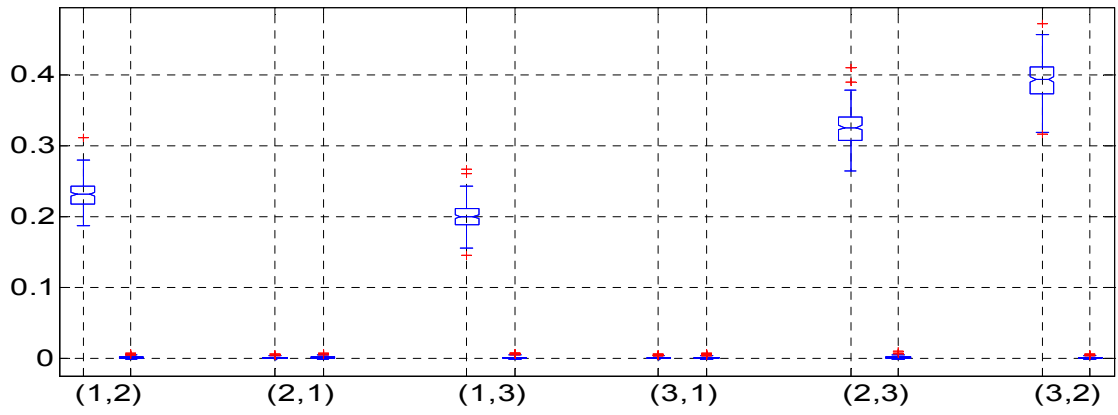
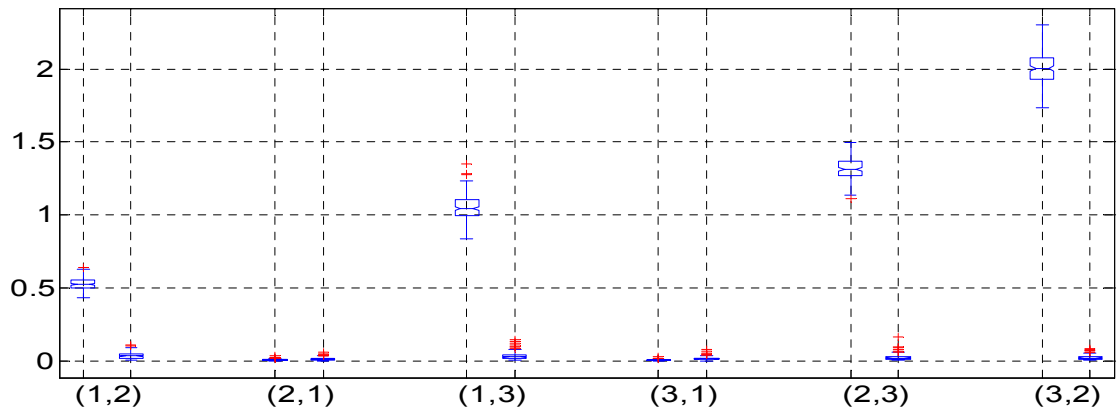


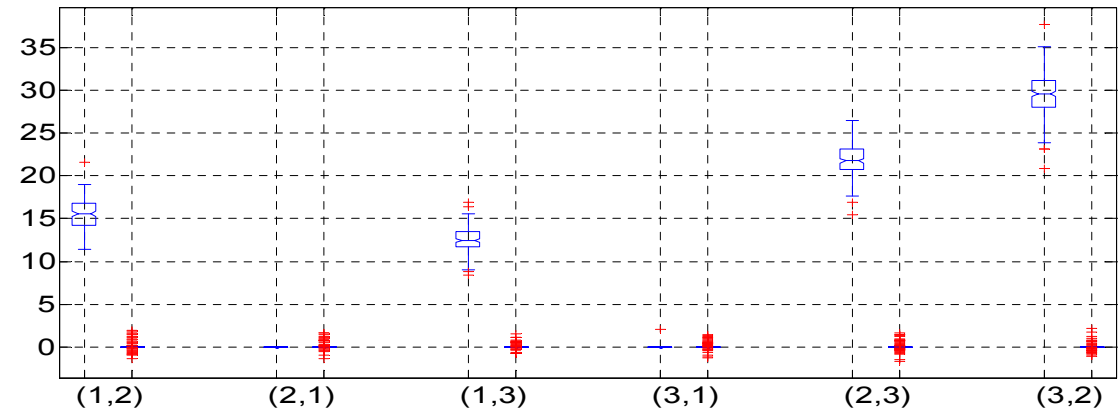
Fig. 5.1 – Boîtes à moustaches des indices (a) WGCI-C, (b) CI-PDC et (c) CTE pour le modèle 4 linéaire ($\alpha = 0.5$ et $\beta = 0.5$) en abscisses : (i, j) représente le flux d'information du signal i vers le signal j , sous les hypothèses H_0 (à droite) et \bar{H}_0 (à gauche) en ordonnées : valeurs de l'indice sous l'hypothèse d'indépendance H_0 (à droite) et sous l'hypothèse \bar{H}_0 (à gauche) pour un flux (i, j) .



(a)

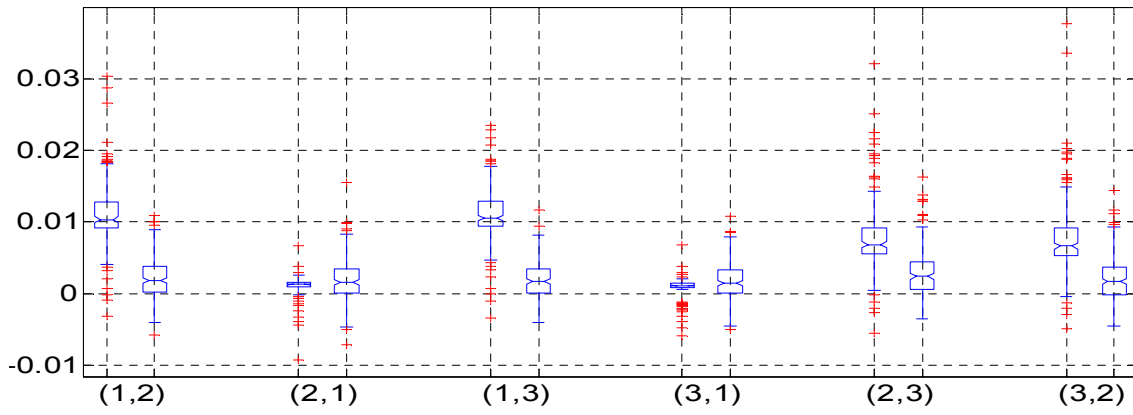


(b)

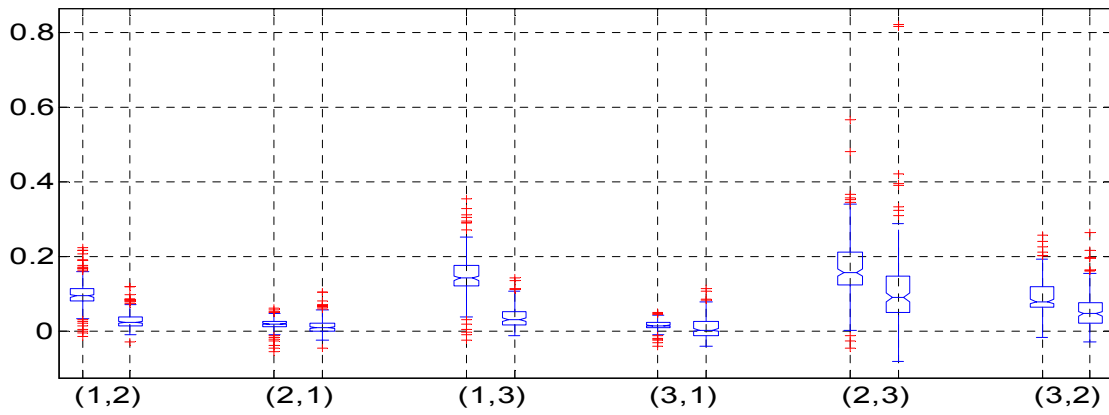


(c)

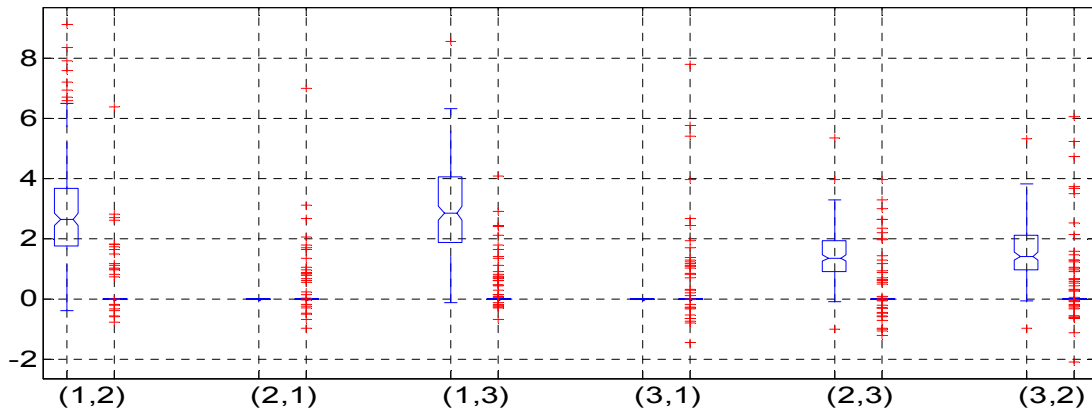
Fig. 5.2 – Boîtes à moustaches des indices (a) WGCI-C, (b) CI-PDC et (c) CTE pour le modèle 4 non linéaire ($\alpha = 0.5$ et $\beta = 0.5$) en abscisses : (i, j) représente le flux d'information du signal i vers le signal j , sous les hypothèses H_0 (à droite) et \bar{H}_0 (à gauche) en ordonnées : résultats sous l'hypothèse d'indépendance H_0 (à droite) et sous l'hypothèse \bar{H}_0 (à gauche) pour un flux (i, j) .



(a)



(b)



(c)

Fig. 5.3 – Boîtes à moustaches pour les indices (a) WGCI-C, (b) CI-PDC et (c) CTE
pour le modèle physiologique
en abscisses : (i, j) représente le flux d'information du signal i vers le signal j ,
sous les hypothèses H_0 (à droite) et \bar{H}_0 (à gauche)
en ordonnées : résultats sous l'hypothèse d'indépendance H_0 (à droite)
et sous l'hypothèse \bar{H}_0 (à gauche) pour un flux (i, j) .

5.3. Analyse des résultats sur modèles physiologiques

Pour le modèle physiologique testé (détaillé dans la section 5.3 du chapitre 5 de la version anglaise), les indices apparaissant comme les plus pertinents en simulation AR vectorielle (linéaire ou non) continuent à bien se comporter et leurs performances peuvent s'apprécier sur la Fig. 5.3. Il convient de constater sur cette figure que les hypothèses H_0 et \bar{H}_0 se distinguent assez nettement pour toutes les approches pour les relations unidirectionnelles $1 \rightarrow 2$ et $1 \rightarrow 3$. Les résultats sont moins nets sur la relation bidirectionnelle $2 \rightleftharpoons 3$. Si l'on fait abstraction des données aberrantes, la méthode CTE apparaît performante, ceci dû en partie à une optimisation de l'estimation de l'ordre des modèles sous les deux hypothèses testées suivant le critère gBIC introduit. Au vu de ces résultats, on peut penser que l'indice WGCI-C pourrait voir ses performances améliorées en considérant une optimisation de l'ordre (ce que nous n'avons pas envisagé jusqu'à présent). Il semble que l'indice issu de la pente de phase perde légèrement de son efficacité sur la situation présente. Toutefois, cette figure met en exergue une certaine dispersion des valeurs quelle que soit la mesure. Aussi, afin de mieux apprécier les performances de nos approches, proposons-nous sur la Fig. 5.4 une représentation sous forme de courbes COR (Caractéristiques Opérationnelles de Réception). Pour les relations unidirectionnelles, *i.e.* $1 \rightarrow 2$ et $1 \rightarrow 3$, les résultats obtenus par les 3 mesures sont presque parfaits. Pour la relation bilatérale $2 \rightleftharpoons 3$, la méthode CTE se révèle très performante et supérieure aux deux autres, la méthode CI-PDC apparaissant comme la plus fragile. Notons que les courbes correspondant aux relations inexistantes $2 \rightarrow 1$ et $3 \rightarrow 1$ sont proches de la diagonale, ce qui est cohérent puisque les deux distributions expérimentales correspondent alors à la même hypothèse H_0 et, donc, se recouvrent.

5.4. Analyse des résultats sur signaux réels

Pour les signaux réels, nous disposons d'enregistrements sur une phase d'installation d'activité rapide de début de crise. Cette phase est caractérisée par un pattern particulier qui trouve son origine dans la région entorhinal/hippocampique et se propage secondairement au cortex périrhinal [Uva 2005]. Ces signaux sont testés par les méthodes WGCI-C, CI-PDC, et CTE. Le signal 1 est enregistré dans l'hippocampe, le signal 2 dans le cortex entorhinal médian et le signal 3 dans le cortex périrhinal. La durée de cette phase d'activité que nous avons considérée est de 10 secondes. Les estimations des indices sont faites sur des fenêtres de 4 secondes avec un pas de 0,125 seconde (49 fenêtres au total). Sur la Fig. 5.5, nous avons représenté l'évolution des différents indices au cours du temps sur la période analysée. La signification des valeurs d'indices révélant l'influence du signal i sur le signal j , et *vice versa*, et de leurs variations peut être appréhendée relativement aux deux boîtes à moustaches représentées sur cette figure respectivement en bas à gauche pour la relation de i vers j et en bas à droite pour la relation de j vers i , boîtes qui ont été élaborées à partir des valeurs d'indices obtenues avec la méthode de brouillage de phase pour simuler l'hypothèse H_0 . Plus précisément, les 10 secondes d'observation ont été recouvertes avec 4 fenêtres de 4 secondes chacune, et se chevauchant 2 à 2 sur 2 secondes. Sur chacune d'elles, 50 réalisations par brouillage de phase ont été obtenues. Les 200 valeurs (4 x 50) de l'indice étudié pour le lien choisi ont ensuite

été fusionnées dans une même boîte à moustaches. Les résultats obtenus par les trois techniques traduisent les mêmes tendances. Les signaux 1 et 2 semblent s'influencer mutuellement sur cette séquence temporelle même si l'on note une diminution modérée de l'influence du signal 1 sur le signal 2 (plus particulièrement avec les méthodes CI-PDC et WGCI-C) sur la deuxième moitié de l'intervalle. Pour ce qui est des relations entre les signaux 1 et 3, elles s'inversent au cours du temps. La relation du signal 1 vers le signal 3 s'avère élevée au début (de l'ordre de grandeur de la relation qui lie le signal 1 au signal 2) et ce quelle que soit la méthode testée. Pour ce qui est de la relation du signal 3 vers le signal 1, elle est inexistante dans les premières secondes pour toutes les méthodes. Quant à la paire de signaux (2, 3), une relation du signal 2 vers le signal 3 apparaît prépondérante pour toutes les méthodes considérées, essentiellement sur la deuxième moitié de l'intervalle d'analyse. A l'inverse, pour aucun des indices, on ne détecte d'influence du signal 3 sur le signal 2. Ces premiers résultats, *a priori* conformes à ceux décrits dans [Uva 2005], demandent une certaine prudence dans la mesure où ils ne portent que sur un cas d'étude et ne pourront être acceptés qu'après une étude à plus grande échelle sur une base de données plus étendue.

5.5. Discussion

De ces différentes expérimentations, nous pouvons conclure à des comportements cohérents et fiables des indices testés dans des situations classiques de modélisation autorégressive linéaire. Ces comportements se maintiennent dans des situations où les non linéarités restent raisonnables, *i.e.* dans le cas de signaux autorégressifs vectoriels non linéaires présentant des couplages linéaires. Testés sur des modèles physiologiques réalistes, ces indices résistent bien aux non linéarités induites par ces situations plus complexes et permettent encore d'établir des graphes pertinents de connectivité effective. Toutefois, ces expérimentations mettent en évidence la sensibilité des approches entropiques au choix des horizons de prédiction. Finalement, la confrontation de ces mesures sur une situation réelle corrobore ces premières analyses, même si, en pareille situation, la vérité-terrain n'est pas accessible mais des hypothèses peuvent être émises. Notons que dans l'esprit d'une mesure de causalité non linéaire et non paramétrique, il serait tout à fait possible d'envisager une technique de régression non linéaire adaptée à l'approche de Granger, prolongeant ainsi le coefficient de corrélation non linéaire standard [Louis Dorr 2007, Wendling 2010]. Rappelons que ce dernier ne cherche à détecter une éventuelle causalité qu'en testant un lien fonctionnel entre un échantillon passé (plus ou moins retardé) de la voie prédictrice et l'échantillon courant de la voie à prédire, ce qui constitue clairement une approche moins fine que celle de Granger.

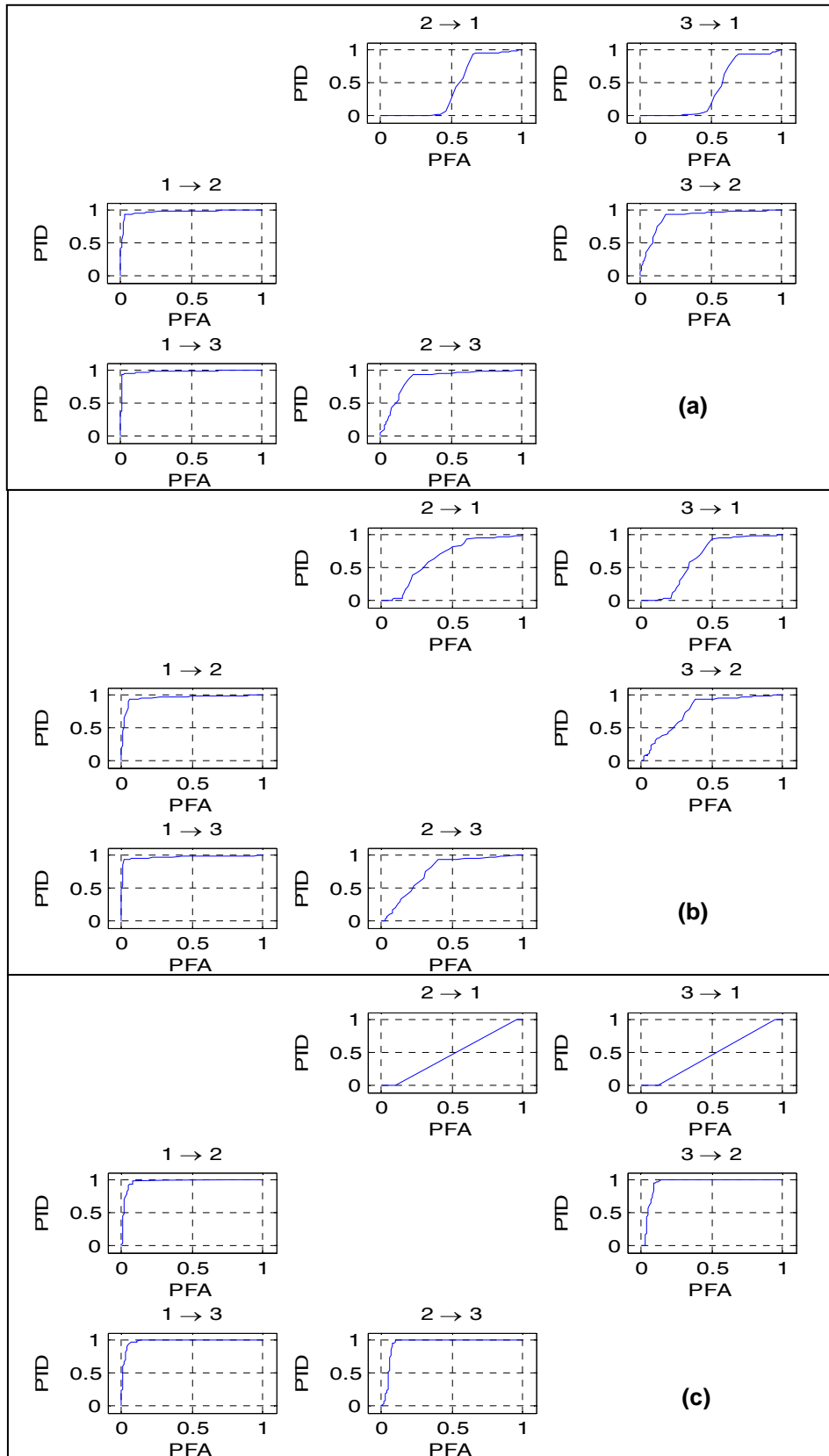
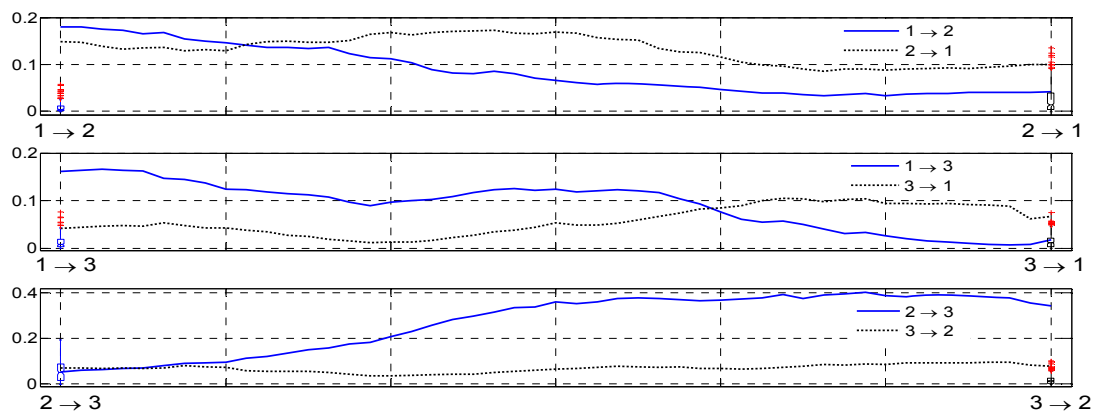
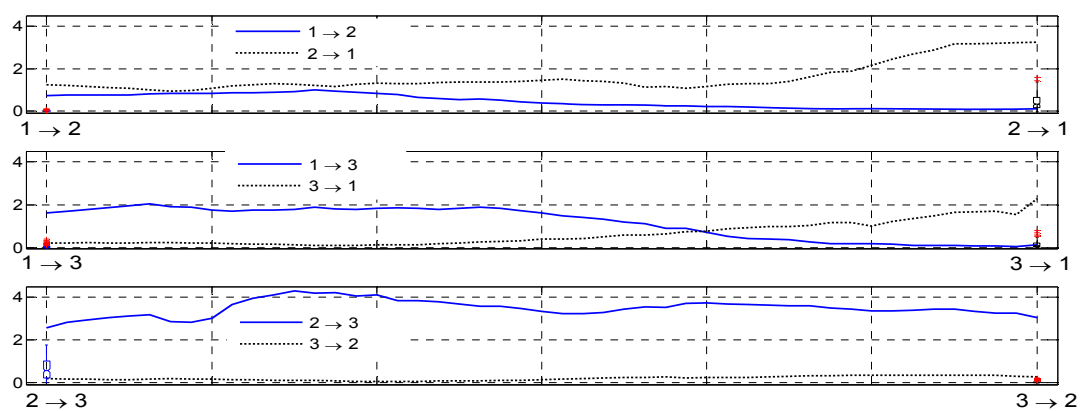


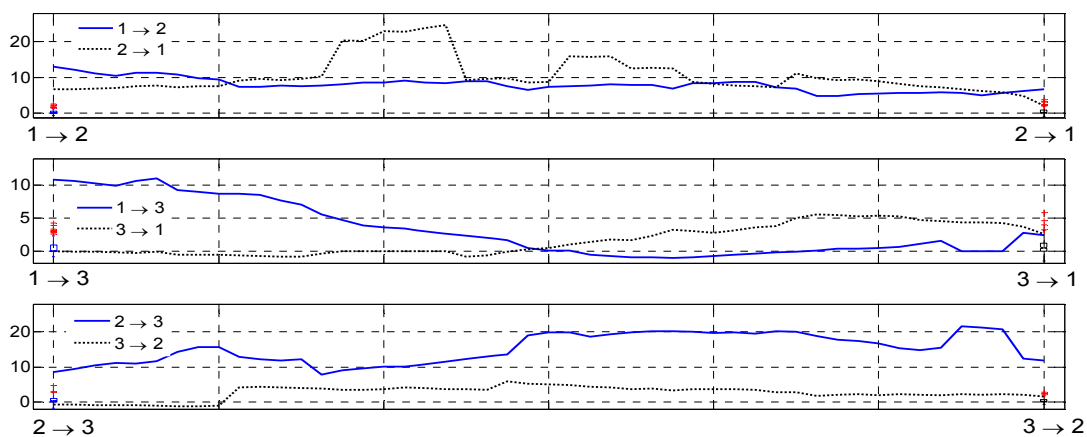
Fig. 5.4 – Courbes COR des méthodes (a) WGCI-C, (b) CI-PDC et (c) CTE pour le modèle physiologique 4
en abscisses : probabilité de fausse alarme (PFA)
en ordonnées : probabilité de bonne détection (PTD).



(a)



(b)



(c)

*Fig. 5.5 – Evolution des indices WGCI-C (a), CI-PDC (b) et CTE (c)
sur une fenêtre temporelle de 10 s des signaux réels
en abscisses : (i, j) représente le flux d'information porté par l'indice testé
du signal i vers le signal j
en ordonnées : courbes d'évolution des indices du signal i vers le signal j et réciproquement
et représentation des boîtes à moustaches sous l'hypothèse d'indépendance H_0
pour les deux directions considérées.*

Conclusion

L'objectif de ce travail était de contribuer à l'analyse de la connectivité cérébrale et plus particulièrement à celle de la connectivité effective, qui, à l'inverse de la connectivité fonctionnelle qu'elle vient enrichir, permet d'attribuer une directivité à la relation entretenue par deux régions cérébrales. Si l'intérêt d'étudier les relations entre les aires du cerveau n'est plus à prouver - les visions "localisationnistes" et "globalistes", qui pouvaient s'opposer jusqu'au siècle dernier quant à rapporter chaque faculté à l'activité d'une structure particulière ou bien au contraire à une activité en réseau, s'accordent aujourd'hui sur une vision "connectiviste" du fonctionnement cérébral [Mangin 1998] - la méthodologie pour y parvenir reste une question ouverte. Bien que cette contribution ait été envisagée essentiellement sur le plan du traitement du signal, un effort a été fait pour décrire le paradigme de la connectivité ainsi que l'approche neurocomputationnelle en neurosciences, afin de mieux situer le travail effectué et ses perspectives.

Dans cette optique, différentes approches ont été envisagées afin d'apporter des éléments de réponse à deux questions essentielles, à savoir l'identification d'influences unilatérales et/ou bilatérales entre structures et la mise en évidence de liens directs et/ou indirects. Notre but était de recenser les principales méthodes proposées dans la littérature, d'en proposer des variantes qui nous semblaient pouvoir être utiles, et de mener une étude expérimentale en simulation pour essayer de dégager des préférences si cela était justifié. Une contrainte que l'on s'était fixée était de ne pas mener cette comparaison uniquement sur des modèles génériques. En effet, la littérature que nous avons consultée et qui concernait ce type d'expérimentation (en lien avec l'étude de connectivité cérébrale) recourait essentiellement à des modèles AR linéaires, des modèles AR non linéaires et des systèmes dynamiques non linéaires ordinairement pertinents dans l'analyse de systèmes chaotiques (systèmes de Lorenz et de Rössler par exemple). Nous avons donc choisi de conserver les premiers (AR linéaires et non linéaires), pour garder un noyau de références expérimentales usuelles, et de préférer aux derniers une simulation de signaux EEG de profondeur (non bruités) au moyen d'un modèle non linéaire physiologiquement réaliste.

Les techniques existantes d'analyse de connectivité que nous avons sélectionnées initialement comprenaient l'indice (linéaire) de causalité de Granger incluant des extensions fréquentielles et/ou conditionnées, et une mesure d'entropie de transfert proposée comme méthode non linéaire et non paramétrique (à des horizons de prédiction près) pour analyser une paire de signaux. Nous avons naturellement considéré cette dernière sous forme conditionnelle pour la détection de liens de

connectivité directs en présence d'un troisième signal. A partir de l'indice de pente de phase introduit par Nolte, nous avons proposé de l'utiliser dans le cadre de la modélisation autorégressive et de compléter la recherche de liens directionnels par une information liée à des temps de propagation éventuels entre groupes neuronaux. Toutes les méthodes requièrent la détermination de N^2 ordres de prédiction, si N est le nombre de voies considérées simultanément, que ce soit pour l'approche de Granger linéaire qui s'appuie sur des représentations AR ou pour les approches entropiques. La sensibilité des indices calculés à ces ordres pouvant être dommageable, ce que nous avons constaté expérimentalement tout particulièrement pour l'entropie de transfert, une analyse de cette sensibilité a été conduite et une stratégie efficiente pour réduire le temps de calcul de la détermination des ordres a été proposée.

Ainsi calibrées, toutes ces méthodes se sont montrées performantes, étant entendu (i) que celles dites "partielles" ou "conditionnelles" (*i.e.* conditionnelles à toute information disponible autre que celle provenant des deux signaux pour lesquels la connectivité est testée) se sont naturellement révélées les seules capables de faire la part entre connectivités directe et indirecte, et (ii) que les signaux étaient modélisés sans bruit additif perturbateur comme on le considère assez naturellement dans le cas d'enregistrements intracérébraux. Les simulations conduites sur signaux AR non linéaires ont confirmé ces premiers résultats obtenus sur modèles linéaires.

L'étude a ensuite été poursuivie sur des signaux modélisés de manière physiologiquement réaliste, au moyen d'un modèle de populations développé antérieurement, pour lequel différents paramétrages ont été recherchés pour à la fois représenter les différents graphes de connectivité déjà utilisés en modélisation AR et simuler des activités épileptogènes. Dans cette étape, seules trois méthodes ont été retenues : l'indice de Granger temporel conditionnel, l'indice de pente de phase basé sur la cohérence dirigée partielle et l'entropie de transfert conditionnée. S'appuyant sur une simulation de l'hypothèse H_0 (absence de connectivité) au moyen de données de substitution (suivant deux options), une analyse statistique des performances de ces trois méthodes a été proposée sous forme de boîtes à moustaches et de courbes COR, non seulement sur le modèle physiologique mais également sur les signaux AR linéaires et non linéaires, pour une seule des quatre topologies (graphes de connectivité) considérées initialement (pour des raisons de temps de calculs). Il s'est avéré que les indicateurs résistent bien aux non linéarités induites par ces situations plus complexes et permettent encore d'établir des graphes cohérents et pertinents de connectivité effective, confirmant nos pressentiments suite aux performances obtenues sur les modèles AR non linéaires.

Une dernière analyse a porté sur des signaux réels enregistrés sur un modèle animal d'épilepsie, dans trois structures – hippocampe, cortex périrhinal, cortex entorhinal médian – lors de l'apparition d'une activité rapide à l'installation de la crise. Bien que l'absence de vérité "terrain" dans ce cas rende plus difficile le jugement, l'utilisation de données de substitution a permis de conjecturer un bon comportement des méthodes, et, globalement, pour les trois méthodes testées, les résultats sont apparus consistants.

Quelles sont les principales conclusions et perspectives de ce travail ? La question posée au début de cette étude était, après avoir constaté les enjeux dans le champ de la connectivité cérébrale, que ce soit dans l'étude de phénomènes cognitifs ou pour les épilepsies chez l'homme, de choisir et éventuellement aménager, sur la base de la littérature et de tests de simulations réalistes, un noyau de méthodes pouvant assurer un compromis entre performances et complexité. Nous pensons avoir contribué à répondre à cette question, aussi bien par la voie bibliographique qu'expérimentale, et pouvoir résumer notre apport en soulignant quelques difficultés : si l'approche non linéaire d'entropie de transfert est séduisante, sa sensibilité aux paramètres d'ordre se traduit par un coût de calculs élevé pour une estimation fiable. D'autre part, l'estimation des distributions de probabilité en dimension élevée peut être confrontée à des problèmes de gestion de mémoire. Comme il pourrait être également intéressant de mesurer une connectivité sur une durée notablement plus longue, les activités transitoires critiques en épilepsie ou les réponses cognitives n'étant pas les seuls phénomènes d'intérêt (on peut penser aux problèmes de prédiction ou de plasticité), il serait donc pertinent de chercher à lever ces verrous, par exemple en utilisant un processeur graphique (GPU : Graphics Processing Unit). Sur un plan plus fondamental, il est possible de reconsidérer les méthodes d'estimation d'entropie de transfert, ou du moins de vérifier plus avant si certaines permettraient des progrès notables. Une autre difficulté, également liée aux déterminations d'ordre, est l'hypothèse markovienne sous-jacente dans l'approche de Granger. Il est patent que les sorties de nos modèles physiologiques correspondent plus à des processus de Markov cachés ou ARMA et le principe des mesures de causalité entre les signaux ainsi produits demanderait une réflexion. La sensibilité des méthodes à des signaux perturbateurs additifs, intervenant avec plus ou moins d'effet même sur des capteurs de champ en profondeur, devrait également être examinée. Enfin, concernant les signaux réels, reste à tester sur une base de données à plus grande échelle (modèle animal par exemple), en croisant avec l'expertise des expérimentateurs, si l'application de ces méthodes peut significativement permettre de vérifier/proposer des hypothèses physiologiques essentielles.

A ce stade de notre recherche, même si nous nous sommes heurtés à des questions restées sans réponse, nous pensons avoir apporté, de manière directe ou indirecte (au sens propre comme au figuré), une pierre à l'édifice des connaissances sur la connectivité cérébrale. L'interaction entre des concepts qui s'opposent et s'affrontent parfois – ségrégation et intégration neuronale, influence et réciprocité, action directe et médiate – dans les réseaux cérébraux produit une information très diversifiée, générant des modèles fortement complexes, aujourd'hui encore non complètement dominés.

Au-delà des perspectives évoquées, une meilleure compréhension de l'organisation des réseaux épileptogènes pourra passer par la considération de techniques multi-modales d'acquisition des signaux dans un souci d'optimisation des résolutions spatiale et temporelle, la finalité thérapeutique étant une délimitation plus fine du volume cérébral à exciser pour supprimer l'activité épileptique, ou du moins la bloquer, ou la faire avorter.

Bibliographie

- [Akaike 1973] Akaike, H., "Information theory and an extension of the maximum likelihood principle," *the Second International Symposium on Information Theory, Budapest, Hungary*. pp. 267-281, 2-8 September 1973.
- [Baccalá 2001] Baccalá, L. A., and Sameshima, K., "Partial directed coherence: a new concept in neural structure determination," *Biological Cybernetics*, vol. 84, no. 6, pp. 463-474, 2001.
- [Bestmann 2008] Bestmann, S., Ruff, C. C., Blankenburg, F., Weiskopf, N., Driver, J., and Rothwell, J. C., "Mapping causal interregional influences with concurrent TMS-fMRI," *Experimental Brain Research*, vol. 191, no. 4, pp. 383-402, 2008.
- [Destexhe 1996] Destexhe, A., Bal, T., McCormick, D. A., and Sejnowski, T. J., "Ionic mechanisms underlying synchronized oscillations and propagating waves in a model of ferret thalamic slices," *Journal of Neurophysiology*, vol. 76, no. 3, pp. 2049-2070, 1996.
- [Edeline 1999] Edeline, J. M., "Learning-induced physiological plasticity in the thalamo-cortical sensory systems: a critical evaluation of receptive field plasticity, map changes and their potential mechanisms," *Progress in Neurobiology*, vol. 57, no. 2, pp. 165-224, 1999.
- [Ermentrout 1998] Ermentrout, B., "Neural networks as spatio-temporal pattern-forming systems," *Reports on Progress in Physics*, vol. 61, no. 4, pp. 353-430, 1998.
- [Faugeras 2009] Faugeras, O., Touboul, J., and Cessac, B., "A constructive mean-field analysis of multi-population neural networks with random synaptic weights and stochastic inputs," *Frontiers in Computational Neuroscience*, vol. 3, no. 1, pp. 1-28, 2009.
- [Ferrerri 2011] Ferreri, F., Pasqualetti, P., Määttä, S., Ponzo, D., Ferrarelli, F., Tononi, G., Mervaala, E., Miniussi, C., and Rossini, P. M., "Human brain connectivity during single and paired pulse transcranial magnetic stimulation," *NeuroImage*, vol. 54, no. 1, pp. 90-102, 2011.
- [Fraser 1986] Fraser, A. M., and Swinney, H. L., "Independent coordinates for strange attractors from mutual information," *Physical Review A*, vol. 33, no. 2, pp. 1134, 1986.
- [Freeman 1973] Freeman, W. J., "A model of the olfactory system," *Neural Modeling*, vol. 1, pp. 41-62, 1973.
- [Freeman 1975] Freeman, W. J., *Mass Action in the Nervous System*, New York: Academic Press, 1975.
- [Freeman 1987] Freeman, W. J., "Simulation of chaotic EEG patterns with a dynamic model of the olfactory system," *Biological Cybernetics*, vol. 56, no. 2, pp. 139-150, 1987.

- [Friston 1994] Friston, K. J., "Functional and effective connectivity in neuroimaging: a synthesis," *Human Brain Mapping*, vol. 2, no. 1-2, pp. 56-78, 1994.
- [Friston 2003] Friston, K. J., Harrison, L., and Penny, W., "Dynamic causal modelling," *NeuroImage*, vol. 19, no. 4, pp. 1273-1302, 2003.
- [Friston 2009] Friston, K. J., "Modalities, modes, and models in functional neuroimaging," *Science*, vol. 326, no. 5951, pp. 399-403, 2009.
- [Frogerais 2008] Frogerais, P., "Modélisation et identification en épilepsie : De la dynamique des populations neuronales aux signaux EEG," LTSI - Laboratoire Traitement du Signal et de l'Image, Université Rennes 1, Rennes, France, 2008.
- [Geweke 1982] Geweke, J. F., "Measurement of linear dependence and feedback between multiple time series," *Journal of the American Statistical Association*, vol. 77, no. 378, pp. 304-313, 1982.
- [Granger 1969] Granger, C. W. J., "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*, vol. 37, no. 3, pp. 424-438, Aug, 1969.
- [Hallett 2000] Hallett, M., "Transcranial magnetic stimulation and the human brain," *Nature*, vol. 406, no. 6792, pp. 147-150, 2000.
- [Honey 2007] Honey, C. J., Kötter, R., Breakspear, M., and Sporns, O., "Network structure of cerebral cortex shapes functional connectivity on multiple time scales," *Proceedings of the National Academy of Sciences*, vol. 104, no. 24, pp. 10240-10245, 2007.
- [Horwitz 2003] Horwitz, B., "The elusive concept of brain connectivity," *NeuroImage*, vol. 19, no. 2, pp. 466-470, 2003.
- [Hubel 1963] Hubel, D. H., and Wiesel, T. N., "Shape and arrangement of columns in cat's striate cortex," *The Journal of Physiology*, vol. 165, no. 3, pp. 559-568, 1963.
- [Hubel 1965] Hubel, D. H., and Wiesel, T. N., "Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat," *Journal of Neurophysiology*, vol. 28, no. 2, pp. 229-289, 1965.
- [Hubel 1968] Hubel, D. H., and Wiesel, T. N., "Receptive fields and functional architecture of monkey striate cortex," *The Journal of Physiology*, vol. 195, no. 1, pp. 215-243, 1968.
- [Jansen 1993] Jansen, B. H., Zouridakis, G., and Brandt, M. E., "A neurophysiologically-based mathematical model of flash visual evoked potentials," *Biological Cybernetics*, vol. 68, no. 3, pp. 275-283, 1993.
- [Jansen 1995] Jansen, B. H., and Rit, V. G., "Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns," *Biological Cybernetics*, vol. 73, no. 4, pp. 357-366, 1995.
- [Kamiński 1991] Kamiński, M. J., and Blinowska, K. J., "A new method of the description of the information flow in the brain structures," *Biological Cybernetics*, vol. 65, no. 3, pp. 203-210, 1991.
- [Kramer 2005] Kramer, M. A., Kirsch, H. E., and Szeri, A. J., "Pathological pattern formation and cortical propagation of epileptic seizures," *Journal of the Royal Society Interface*, vol. 2, no. 2, pp. 113-127, 2005.

- [Kramer 2006] Kramer, M. A., Lopour, B. A., Kirsch, H. E., and Szeri, A. J., "Bifurcation control of a seizing human cortex," *Physical Review E*, vol. 73, no. 4, pp. 041928-1:041928-16, 2006.
- [Lopes Da Silva 1974] Lopes Da Silva, F. H., Hoeks, A., Smits, H., and Zetterberg, L. H., "Model of brain rhythmic activity," *Biological Cybernetics*, vol. 15, no. 1, pp. 27-37, 1974.
- [Lopes Da Silva 1976] Lopes Da Silva, F. H., Van Rotterdam, A., Barts, P., Van Heusden, E., and Burr, W., *Model of Neuronal Populations - The Basic Mechanism of Rhythmicity*, Amsterdam: Elsevier, 1976.
- [Lopes Da Silva 1989] Lopes Da Silva, F. H., Pijn, J. P., and Boeijinga, P., "Interdependence of EEG signals: linear vs. nonlinear associations and the significance of time delays and phase shifts," *Brain Topography*, vol. 2, no. 1, pp. 9-18, 1989.
- [Lopes Da Silva 2003] Lopes Da Silva, F. H., Blanes, W., Kalitzin, S. N., Parra, J., Suffczynski, P., and Velis, D. N., "Dynamical diseases of brain systems: different routes to epileptic seizures," *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 5, pp. 540-548, 2003.
- [Lorenz 1963] Lorenz, E. N., "Deterministic Nonperiodic Flow," *Journal of the Atmospheric Sciences*, vol. 20, pp. 130-141, 1963.
- [Louis Dorr 2007] Louis Dorr, V., Caparos, M., Wendling, F., Vignal, J. P., and Wolf, D., "Extraction of reproducible seizure patterns based on EEG scalp correlations," *Biomedical Signal Processing and Control*, vol. 2, no. 3, pp. 154-162, 2007.
- [Mangin 1998] Mangin, J. F., Régis, J., Poline, J. B., Rivière, D., Poupon, C., Poupon, F., Papadopoulos, D., Delaye, F., Pizzato, O., and Coulon, O., "Place de l'anatomie dans la cartographie fonctionnelle du cerveau," *Annales de l'Institut Pasteur*, vol. 9, no. 3, pp. 243-258, 1998.
- [Martinerie 1992] Martinerie, J. M., Albano, A. M., Mees, A. I., and Rapp, P. E., "Mutual information, strange attractors, and the optimal estimation of dimension," *Physical Review A*, vol. 45, no. 10, pp. 7058-7064, 1992.
- [Massimini 2005] Massimini, M., Ferrarelli, F., Huber, R., Esser, S. K., Singh, H., and Tononi, G., "Breakdown of cortical effective connectivity during sleep," *Science*, vol. 309, no. 5744, pp. 2228, 2005.
- [McIntosh 1994] McIntosh, A. R., and Gonzalez-Lima, F., "Structural equation modeling and its application to network analysis in functional brain imaging," *Human Brain Mapping*, vol. 2, no. 1-2, pp. 2-22, 1994.
- [Mountcastle 1957] Mountcastle, V. B., "Modality and topographic properties of single neurons of cat's somatic sensory cortex," *Journal of Neurophysiology*, vol. 20, no. 4, pp. 408-434, 1957.
- [Nolte 2008] Nolte, G., Ziehe, A., Nikulin, V. V., Schlögl, A., Krämer, N., Brismar, T., and Müller, K. R., "Robustly estimating the flow direction of information in complex physical systems," *Physical Review Letters*, vol. 100, no. 23, pp. 234101, June, 2008.
- [Ooyen 2001] Ooyen, A. V., "Competition in the development of nerve connections: a review of models," *Network: Computation in Neural Systems*, vol. 12, no. 1, pp. 1-47, 2001.
- [Paus 1997] Paus, T., Jech, R., Thompson, C. J., Comeau, R., Peters, T., and Evans, A. C., "Transcranial magnetic stimulation during positron emission tomography: a new

method for studying connectivity of the human cerebral cortex," *Journal of Neuroscience*, vol. 17, no. 9, pp. 3178-3184, 1997.

- [Ruff 2009] Ruff, C. C., Driver, J., and Bestmann, S., "Combining TMS and fMRI: From 'virtual lesions' to functional-network accounts of cognition," *Cortex*, vol. 45, no. 9, pp. 1043-1049, 2009.
- [Sabesan 2007] Sabesan, S., Narayanan, K., Prasad, A., Iasemidis, L. D., Spanias, A., and Tsakalis, K., *Information flow in coupled nonlinear systems: application to the epileptic human brain*, New York: Springer, 2007.
- [Sabesan 2009a] Sabesan, S., Chakravarthy, N., Tsakalis, K., Pardalos, P., and Iasemidis, L., "Measuring resetting of brain dynamics at epileptic seizures: application of global optimization and spatial synchronization techniques," *Journal of Combinatorial Optimization*, vol. 17, no. 1, pp. 74-97, 2009a.
- [Sabesan 2009b] Sabesan, S., Good, L., Tsakalis, K., Spanias, A., Treiman, D., and Iasemidis, L., "Information flow and application to epileptogenic focus localization from intracranial EEG," *IEEE Transactions on Neural Systems and Rehabilitation Engineering: a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 17, no. 3, pp. 244-253, June, 2009b.
- [Saito 1981] Saito, Y., and Harashima, H., *Tracking of information within multichannel EEG record-causal analysis in EEG*, Amsterdam: Elsevier-North-Holland Biomedical Press, 1981.
- [Sakkalis 2011] Sakkalis, V., "Review of advanced techniques for the estimation of brain connectivity measured with EEG/MEG," *Computers in Biology and Medicine*, vol. 41, no. 12, pp. 1110-1117, 2011.
- [Schreiber 2000] Schreiber, T., "Measuring information transfer," *Physical Review Letters*, vol. 85, no. 2, pp. 461-464, 2000.
- [Schwarz 1978] Schwarz, G., "Estimating the dimension of a model," *The Annals of Statistics*, vol. 6, no. 2, pp. 461-464, 1978.
- [Skinner 2005a] Skinner, F. K., Bazzazi, H., and Campbell, S. A., "Two-cell to N-cell heterogeneous, inhibitory networks: Precise linking of multistable and coherent properties," *Journal of Computational Neuroscience*, vol. 18, no. 3, pp. 343-352, 2005a.
- [Skinner 2005b] Skinner, F. K., Chung, J. Y. J., Ncube, I., Murray, P. A., and Campbell, S. A., "Using heterogeneity to predict inhibitory network model characteristics," *Journal of Neurophysiology*, vol. 93, no. 4, pp. 1898-1907, 2005b.
- [Sporns 2000] Sporns, O., Tononi, G., and Edelman, G. M., "Connectivity and complexity: the relationship between neuroanatomy and brain dynamics," *Neural Networks*, vol. 13, no. 8-9, pp. 909-922, 2000.
- [Sporns 2004] Sporns, O., Chialvo, D. R., Kaiser, M., and Hilgetag, C. C., "Organization, development and function of complex brain networks," *Trends in Cognitive Sciences*, vol. 8, no. 9, pp. 418-425, 2004.
- [Sporns 2005] Sporns, O., Tononi, G., and Kötter, R., "The human connectome: a structural description of the human brain," *PLoS Computational Biology*, vol. 1, no. 4, pp. 245-251, 2005.
- [Sporns 2007] Sporns, O., 2007, http://www.scholarpedia.org/article/Brain_connectivity.

- [Sporns 2010] Sporns, O., *Networks of the Brain*, London, England: The MIT Press, 2010.
- [Suffczynski 2001] Suffczynski, P., Kalitzin, S., Pfurtscheller, G., and Lopes Da Silva, F. H., "Computational model of thalamo-cortical networks: dynamical control of alpha rhythms in relation to focal attention," *International Journal of Psychophysiology*, vol. 43, no. 1, pp. 25-40, 2001.
- [Suffczynski 2004] Suffczynski, P., Kalitzin, S., and Lopes Da Silva, F. H., "Dynamics of non-convulsive epileptic phenomena modeled by a bistable neuronal network," *Neuroscience*, vol. 126, no. 2, pp. 467-484, 2004.
- [Suffczynski 2006] Suffczynski, P., Lopes Da Silva, F. H., Parra, J., Velis, D. N., Bouwman, B. M., Van Rijn, C. M., Van Hese, P., Boon, P., Khosravani, H., and Derchansky, M., "Dynamics of epileptic phenomena determined from statistics of ictal transitions," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 3, pp. 524-532, 2006.
- [Tononi 1994] Tononi, G., Sporns, O., and Edelman, G. M., "A measure for brain complexity: relating functional segregation and integration in the nervous system," *Proceedings of the National Academy of Sciences*, vol. 91, no. 11, pp. 5033-5037, 1994.
- [Tononi 2003] Tononi, G., and Sporns, O., "Measuring information integration," *BMC Neuroscience*, vol. 4, no. 1, pp. 31-50, 2003.
- [Touboul 2011] Touboul, J., Wendling, F., Chauvel, P., and Faugeras, O., "Neural mass activity, bifurcations, and epilepsy," *Neural Computation*, vol. 23, no. 12, pp. 3232-3286, 2011.
- [Traub 1982] Traub, R. D., and Wong, R. K., "Cellular mechanism of neuronal synchronization in epilepsy," *Science*, vol. 216, no. 4547, pp. 745-747, 1982.
- [Traub 2001] Traub, R. D., Whittington, M. A., Buhl, E. H., Lebeau, F. E. N., Bibbig, A., Boyd, S., Cross, H., and Baldeweg, T., "A possible role for gap junctions in generation of very fast EEG oscillations preceding the onset of, and perhaps initiating, seizures," *Epilepsia*, vol. 42, no. 2, pp. 153-170, 2001.
- [Traub 2003] Traub, R. D., Pais, I., Bibbig, A., Lebeau, F. E. N., Buhl, E. H., Hormuzdi, S. G., Monyer, H., and Whittington, M. A., "Contrasting roles of axonal (pyramidal cell) and dendritic (interneuron) electrical coupling in the generation of neuronal network oscillations," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 3, pp. 1370-1374, 2003.
- [Traub 2005] Traub, R. D., Contreras, D., Cunningham, M. O., Murray, H., Lebeau, F. E. N., Roopun, A., Bibbig, A., Wilent, W. B., Higley, M. J., and Whittington, M. A., "Single-column thalamocortical network model exhibiting gamma oscillations, sleep spindles, and epileptogenic bursts," *Journal of Neurophysiology*, vol. 93, no. 4, pp. 2194-2232, 2005.
- [Treserras 2008] Treserras, S., "Etudes sur la connectivité cérébrale: aspects méthodologiques et applications au cerveau au repos, à la motricité et à la lecture," *Imagerie Médicale et Handicaps Neurologiques*, INSERM UMR 825, Université de Toulouse, Université Toulouse III-Paul Sabatier, Toulouse, 2008.
- [Uva 2005] Uva, L., Librizzi, L., Wendling, F., and De Curtis, M., "Propagation dynamics of epileptiform activity acutely induced by Bicuculline in the hippocampal-parahippocampal region of the isolated guinea pig brain," *Epilepsia*, vol. 46, no. 12, pp. 1914-1925, 2005.

- [Valdes-Sosa 2011] Valdes-Sosa, P. A., Roebroek, A., Daunizeau, J., and Friston, K., "Effective connectivity: Influence, causality and biophysical modeling," *NeuroImage*, vol. 58, no. 2, pp. 339-361, 2011.
- [Van Drongelen 2005] Van Drongelen, W., Lee, H. C., Hereld, M., Chen, Z., Elsen, F. P., and Stevens, R. L., "Emergent epileptiform activity in neural networks with weak excitatory synapses," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 13, no. 2, pp. 236-241, 2005.
- [Van Drongelen 2007] Van Drongelen, W., Lee, H. C., Stevens, R. L., and Hereld, M., "Propagation of seizure-like activity in a model of neocortex," *Journal of Clinical Neurophysiology*, vol. 24, no. 2, pp. 182-188, 2007.
- [Wendling 2000] Wendling, F., Bellanger, J. J., Bartolomei, F., and Chauvel, P., "Relevance of nonlinear lumped-parameter models in the analysis of depth-EEG epileptic signals," *Biological Cybernetics*, vol. 83, no. 4, pp. 367-378, 2000.
- [Wendling 2001] Wendling, F., Bartolomei, F., Bellanger, J. J., and Chauvel, P., "Interpretation of interdependencies in epileptic signals using a macroscopic physiological model of the EEG," *Clinical Neurophysiology*, vol. 112, no. 7, pp. 1201-1218, 2001.
- [Wendling 2002] Wendling, F., Bartolomei, F., Bellanger, J. J., and Chauvel, P., "Epileptic fast activity can be explained by a model of impaired GABAergic dendritic inhibition," *European Journal of Neuroscience*, vol. 15, no. 9, pp. 1499-1508, 2002.
- [Wendling 2005] Wendling, F., Hernandez, A., Bellanger, J. J., Chauvel, P., and Bartolomei, F., "Interictal to ictal transition in human temporal lobe epilepsy: insights from a computational model of intracerebral EEG," *Journal of Clinical Neurophysiology*, vol. 22, no. 5, pp. 343-356, 2005.
- [Wendling 2010] Wendling, F., Chauvel, P., Biraben, A., and Bartolomei, F., "From Intracerebral EEG Signals to Brain Connectivity: Identification of Epileptogenic Networks in Partial Epilepsy," *Frontiers in Systems Neuroscience*, vol. 4, no. 154, pp. 1-13, 2010.
- [Wiener 1956] Wiener, N., *The Theory of Prediction*, New York: McGraw-Hill, 1956.
- [Wilson 1972] Wilson, H. R., and Cowan, J. D., "Excitatory and inhibitory interactions in localized populations of model neurons," *Biophysical Journal*, vol. 12, no. 1, pp. 1-24, 1972.
- [Wilson 1973] Wilson, H. R., and Cowan, J. D., "A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue," *Biological Cybernetics*, vol. 13, no. 2, pp. 55-80, 1973.
- [Yang 2010] Yang, C., Le Bouquin Jeannès, R., and Faucon, G., "Determining the flow direction of causal interdependence in multivariate time series," *18th European Signal Processing Conference (EUSIPCO), Aalborg, Denmark*, pp. 636-640, Aug. 23-27 2010.
- [Yang 2011] Yang, C., Le Bouquin Jeannès, R., Faucon, G., and Shu, H., "Extracting information on flow direction in multivariate time series," *IEEE Signal Processing Letters*, vol. 18, no. 4, pp. 251-254, April 2011.
- [Yang 2012] Yang, C., Le Bouquin Jeannès, R., Faucon, G., and Shu, H., "Detecting information flow direction in multivariate linear and nonlinear models," *Signal Processing*, 2012.
- [Zetterberg 1978] Zetterberg, L. H., Kristiansson, L., and Mossberg, K., "Performance of a model for a local neuron population," *Biological Cybernetics*, vol. 31, no. 1, pp. 15-26, 1978.

**Contribution to
effective connectivity
analysis in epilepsy**

Acknowledgements

This thesis was performed in the frame of CRIBs (Centre de Recherche en Information Biomédicale sino-français), an international associate French-Chinese laboratory (Université de Rennes 1 - France, INSERM - France, SouthEast University – China).

First of all, I would like to thank my supervisor, Professor Régine LE BOUQUIN JEANNES. Her rigorous and meticulous attitude of research, her selfless work passion, her absorption in her hard work and her exceptionally vigorous and resolute work style gave me a deep impression and will remain a model for me and an example to follow in the future. I also give my great gratitude to Associate Professor Jean-Jacques BELLANGER for his fruitful discussions and help in developing methods, and his contribution in the preparation of the defense. I should thank Régine and Jean-Jacques again because of their help and the time they spent on my thesis even during nights and weekends. I would also like to express my heartfelt gratitude to Professor Gérard FAUCON for his support when I started my research and for the discussions we had on some approaches.

I would like to thank Professor Limin LUO who firstly accepted me as his doctor student in the Laboratory of Image and Science Technology (LIST) of Southeast University and then gave me the chance to do my PHD in France. I also should give my great appreciation to Professor Huazhong SHU for his help during my whole thesis and his useful suggestions. I also felt grateful to Professor Lotfi SENHADJI who accepted me in his laboratory (LTSI – Laboratoire Traitement du Signal et de l'Image, Université de Rennes 1). I also thank Fabrice WENDLING for advice and discussions, especially on the physiological model.

I would like to thank Professor Valérie LOUIS DORR and Professor Jacques DUCHENE who accepted to review my document as well as Jean-Marc EDELINE, senior scientist at CNRS, who accepted to preside at the jury.

I cannot forget to thank all the people who gave me help, worked and studied in LTSI between 2008 and 2012 and in LIST between 2007 and 2012.

Otherwise, I appreciated Régine and her family (Eric, Thomas, Juline and Marie), Gérard and his wife (Marie-Jeanne) again for their kindness and gracious hospitality. They made me feel good and warm in a foreign country.

Finally, I should express my deepest gratitude to my wife, Xiaojia WANG, and to our parents (Guochun YANG, Chanjuan DU, Zhiming WANG and Yuefen HUANG) and our

coming baby. When I was disappointed and downhearted, when I had setbacks and difficulties, they always supported me and their love encouraged me to succeed.

Abstract

Our work deals with effective connectivity to detect and quantify relations between cerebral structures involved in the initiation and the diffusion of epileptic seizures, aiming at establishing flow information propagation graphs. Defined as a potential connection to neighbouring or distant entities, connectivity can be derived at a cerebral level according to three modes. Anatomical connectivity refers to physical or structural links that exist between different cortical areas. Functional connectivity characterizes the network activity of the brain. Effective connectivity complements the notions of structural and functional connectivity introducing the concept of causal influence of one neuronal system over another one, either directly or indirectly. In this context, we study different approaches to answer two main questions, (i) the identification of unidirectional and bidirectional relations, (ii) the discrimination between direct and indirect links. Firstly, we investigate the Granger causality index as well as its extended frequential and/or conditional versions. Then, we explore a phase slope index and we introduce a new index based on partial directed coherence to complete the search of directional links providing information on propagation delays between neuronal ensembles. In a third step, we focus on techniques derived from information theory, such as transfer entropy, we selected as a nonlinear and nonparametric method computed from two signals. This method is considered in its conditional form to detect direct links taking into account the presence of a third signal. Since this technique is sensitive to calibration parameters such as the model order, a "greedy" strategy is proposed to optimize the order estimation based on the Bayesian information criterion. All approaches are evaluated and compared using Monte Carlo experiments on linear and nonlinear autoregressive models and also on physiology-based models and real signals recorded on an animal-model (guinea-pig) during a particular phase of a seizure corresponding to a narrowband tonic activity. Results on simulated signals allow us to establish coherent and consistent propagation graphs. For the real signals, without any ground-truth, which makes the assessment difficult, the use of surrogate data allows us to speculate a good behavior of our techniques and, for the three approaches tested, results appear coherent.

Contents

List of abbreviations	1
Introduction	3
Chapter 1. Clinical context and problem statement	7
1.1. Human brain.....	7
1.1.1. Location.....	7
1.1.2. External morphology	7
1.1.3. Internal morphology	9
1.1.3.1. The cerebral cortex or grey matter	9
1.1.3.2. White matter.....	10
1.2. Epilepsy.....	11
1.2.1. Generalities	11
1.2.1.1. Prevalence	11
1.2.1.2. Incidence.....	11
1.2.1.3. Mortality	11
1.2.2. Historical overview	12
1.2.3. Etiology	13
1.3. Towards solutions	15
1.3.1. Scientific and medical advances.....	15
1.3.1.1. Fundamental research	15
1.3.1.2. Clinical research	16
1.3.1.2.1. Pharmacotherapy	16
1.3.1.2.2. Surgical therapy.....	17
1.3.2. Problem statement.....	18
Chapter 2. Connectivity and models	21
2.1. Brain connectivity	21
2.1.1. Generalities	22
2.1.1.1. Modes of brain connectivity	22
2.1.1.2. Graph theory formalism and techniques for brain connectivity.....	23
2.1.2. Observational and computational methods	25
2.1.2.1. Structural connectivity.....	25
2.1.2.2. Functional and effective connectivity	26

2.1.2.2.1.	Overview on instrumental modalities	26
2.1.2.2.2.	Data processing in functional connectivity	28
2.1.2.2.2.1.	Linear techniques	28
2.1.2.2.2.2.	Nonlinear techniques	28
2.1.2.2.3.	Data processing in effective connectivity	29
2.1.2.2.3.1.	Model-based effective connectivity techniques	30
2.1.2.2.3.2.	Data-driven (model free) effective connectivity techniques	31
2.1.3.	Relations between structural, functional and effective connectivity	34
2.2.	Neurocomputational models for epilepsy	35
2.2.1.	Microscopic models	36
2.2.2.	Macroscopic models	37

Chapter 3. Experimental protocol..... 43

3.1.	Database	45
3.1.1.	Linear autoregressive models	45
3.1.2.	Nonlinear autoregressive models	45
3.1.3.	Physiological models	46
3.1.3.1.	First scenario: Model 1	49
3.1.3.2.	Second scenario: Model 2	50
3.1.3.3.	Third scenario: Model 3	50
3.1.3.4.	Fourth scenario: Model 4	50
3.1.4.	Real signals	51
3.2.	Methodology of evaluation	51
3.2.1.	Determination of graphs	51
3.2.2.	Analysis of the distributions under H_0	52

Chapter 4. Methods..... 55

4.1.	Granger causality	56
4.1.1.	Time domain	57
4.1.1.1.	Bivariate case	57
4.1.1.2.	Multivariate case	58
4.1.2.	Frequency domain	59
4.1.2.1.	Bivariate case	59
4.1.2.2.	Trivariate case	62
4.2.	Phase slope index and associated causality indices	66
4.2.1.	Phase slope index	67
4.2.2.	Causality index using partial coherence	70
4.2.3.	Causality index using directed coherence and directed transfer function	71
4.2.4.	Causality index using partial directed coherence	73
4.3.	Transfer entropy and conditional transfer entropy	74
4.3.1.	Transfer entropy	75
4.3.1.1.	Selection of k and l	78
4.3.1.2.	Selection of radius r	79
4.3.2.	Conditional transfer entropy	80
4.3.2.1.	Selection of k , l and s	81
4.3.2.2.	Selection of radius r	82
4.4.	Discussion	82

Chapter 5. Experimental results.....	85
5.1. First selection of techniques from linear AR models.....	85
5.1.1. Granger Causality	86
5.1.1.1. WGCI-P vs WGCI-C	87
5.1.1.2. FGCI-P vs FGCI-C.....	89
5.1.1.3. Discussion and conclusion	90
5.1.2. Phase slope index and causality index.....	92
5.1.2.1. PSI-OC.....	92
5.1.2.2. PSI-OC vs CI-PC	93
5.1.2.3. PSI-OC vs CI-DCOH and CI-DTF.....	94
5.1.2.4. CI-DCOH and CI-DTF vs CI-PDC.....	95
5.1.2.5. Discussion and conclusion	97
5.1.3. Transfer entropy and conditional transfer entropy	97
5.1.3.1. TE.....	98
5.1.3.2. CTE.....	102
5.1.3.3. Discussion and conclusion	107
5.2. Second selection of techniques from nonlinear AR models	108
5.2.1. Granger Causality	108
5.2.1.1. WGCI-C	108
5.2.1.2. FGCI-C.....	110
5.2.1.3. Discussion and conclusion	111
5.2.2. Phase slope index and causality index.....	111
5.2.2.1. CI-DCOH and CI-DTF vs CI-PDC.....	111
5.2.2.2. Discussion and conclusion	112
5.2.3. Transfer entropy and conditional transfer entropy	113
5.2.3.1. TE.....	113
5.2.3.2. CTE.....	114
5.2.3.3. Discussion and conclusion	115
5.3. Validation phase.....	116
5.3.1. Results on the linear model (model 4)	117
5.3.2. Results on the nonlinear model (model 4)	117
5.3.3. Results on the physiological model (model 4)	118
5.4. Tests on real signals	121
5.5. Discussion and conclusion.....	124
Conclusion.....	127
Appendix A. Power spectral densities of the signals	131
Appendix B. Parameters of the physiological models.....	133
B.1. First scenario (model 1)	133
B.2. Second scenario (model 2)	133
B.3. Third scenario (model 3)	134
B.4. Fourth scenario (model 4).....	134
Appendix C. Typology of the methods.....	135
Appendix D. Determining the model order (AIC, BIC).....	137

D.1.	Introduction	137
D.2.	Akaike's information criterion and the bayesian information criterion	137
Appendix E. Notations.....		139
References		141

List of abbreviations

AEDs	AntiEpileptic Drugs
AIC	Akaike's Information Criterion
AR	AutoRegressive
ARX	AutoRegressive exogenous
BIC	Bayesian Information Criterion
BKG	BackGround activity
BOLD	Blood-Oxygen-Level-Dependent
CBF	Cerebral Blood Flow
CI	Causality Index
CI-DCOH	Causality Index using DCOH function
CI-DTF	Causality Index using DTF function
CI-PC	Causality Index using PC function
CI-PDC	Causality Index using PDC function
CNS	Central Nervous System
CTE	Conditional Transfer Entropy
DCM	Dynamic Causal Modeling
DCOH	Directed COHerence
dMRI	diffusion Magnetic Resonance Imaging
DSI	Diffusion Spectrum Imaging
DTF	Directed Transfer Function
DTI	Diffusion Tensor Imaging
DWI	Diffusion Weighted Imaging
ECoG	ElectroCorticoGraphy
EEG	ElectroEncephaloGraphy
EZ	Epileptogenic Zone
FDA	Food and Drug Administration
FFT	Fast Fourier Transform
FGC	Frequency Geweke Causality
FGCI-C	Conditional Geweke Causality Index in the Frequency domain
FGCI-P	Geweke Causality Index in the Frequency domain by Pairwise analysis
fMRI	functional Magnetic Resonance Imaging
FOA	Fast Onset Activity

gAIC	generalized Akaike Information Criterion
gBIC	generalized Bayesian Information Criterion
GPU	Graphics Processing Unit
GS	Generalized Synchronization
IBE	International Bureau for Epilepsy
iEEG	intracerebral EEG
ILAE	International League Against Epilepsy
LFP	Local Field Potential
MEG	MagnetoEncephaloGraphy
MI	Mutual Information
MRI	Magnetic Resonance Imaging
MRS	Magnetic Resonance Spectroscopy
MSC	Magnitude Squared Coherence
MTLE	Mesial Temporal Lobe Epilepsy
OC	Ordinary Coherence
PC	Partial Coherence
PDC	Partial Directed Coherence
PET	Positron Emission Tomography
PLV	Phase Locking Value
PMI	Partial Mutual Information
PS	Phase Synchronization
PSD	Power Spectral Density
PSI	Phase Slope Index
PSI-OC	PSI using the OC function
sd	standard deviation
SDE	Stochastic Differential Equations
SPECT	Single Photon Emission Computed Tomography
TE	Transfer Entropy
TMS	Transcranial Magnetic Stimulation
VAR	Vectorial AutoRegressive
WAGS	Wiener-Akaike-Granger-Schweder
WHO	World Health Organization
WGC	Wiener-Granger Causality
WGCI-C	Conditional Wiener-Granger Causality Index
WGCI-P	Wiener-Granger Causality Index by Pairwise analysis

Introduction

Patients suffering from refractory epilepsy can benefit from a pre-surgical evaluation to help in delimitating the epileptogenic zone (or network) which is responsible for the initiation or propagation of the seizures. It can be done through the recordings of intracerebral signals. Even if this procedure is invasive, it allows a more precise exploration of the brain. Our study aims at extracting characteristics of the epileptogenic network by applying statistical techniques for connectivity analysis.

Possibly defined as a potential connection to neighbouring or distant entities, connectivity can be derived at a cerebral level according to three modes:

- Anatomical (or structural) connectivity: it refers to physical or structural links that exist between different cortical areas. The word "structure" is considered in a wide sense and can represent microscopic elements, *i.e.* at the neuronal scale (synaptic connection between neurons), mesoscopic elements, *i.e.* at the scale of neuronal populations (synaptic ensembles or "bundles" of connections between networks of cortical columns), or macroscopic elements, *i.e.* at the scale of cerebral areas (inter-connection according to different paths between distinct areas of the brain). In essence, the anatomical connectivity underlies the notion of architectural cerebral complexity
- Functional connectivity: this terminology characterizes the network activity of the brain. This concept, which deals with statistics, focuses on the variations of "cooperation" (co-activating structures), and so, on the notion of dependence (including eventually independence of course) between distributed neuronal units (or groups of neurons) that are often spatially distant. The study of functional networks not only includes the study of interactions between these different entities but also the study of the evolution of their interactions across time. Functional connectivity differs from structural connectivity since an anatomical connectivity may, in a non stationary manner, support a functional link. In the same way, functional connectivity does not necessarily imply (direct) anatomical connectivity between these structures since their interdependence may be the result of a third mediating structure, or even a group of other structures. Statistical dependence can be estimated from correlation measures, covariance, spectral coherence, phase synchronization, mutual information. Generally, this connectivity does not refer explicitly to any directivity or causality notion. Only strengths of coupling are taken into account

- Effective connectivity: it complements the notions of structural and functional connectivity introducing the concept of causal influence of one neuronal system over another one, either directly or indirectly. Beyond functional connectivity which reveals a coordinate (non independent) functioning of two systems, the effective connectivity introduces the notion of cause and effect, since it searches for expressing functional dependence of a system under the influence of a second one. This notion refers to that of information flow if we consider the brain as an organ of distributed information processing. Theoretically, causal effects can be inferred through the use of systematic disturbances of the system, or, when these disturbances are impossible to produce, through the prediction analysis given temporal series (since, according to the causality principle, causes must precede their consequences and allow to predict them, at least partially). To investigate this connectivity, some techniques need to specify a model to generate observations that can be summarized in a set of parameters (structural parameters or parameters that modify local temporal dynamics), some others being non parametric.

A connectivity model of the brain can be formally represented with a propagation graph, this graph being weighted and oriented in the case of effective connectivity. Weighting reveals the strength of the connection between the different elements involved in the system. The orientations we look for in effective connectivity are represented by arrows that are unidirectional or bidirectional in case of reciprocal influences (we can imagine that two neuronal populations can stimulate each other in mirror through distinct nervous bundles). Thus, the functional connectivity forms a matrix, in which each element codes a statistical dependence between two elements of the system, and which is symmetric when the indicator of dependence is symmetric itself (as it is the case of the linear correlation for example). Then, this matrix is thresholded to obtain binary non oriented graphs. As for the effective connectivity, it leads more often to a non symmetric matrix, and applying a threshold to this matrix introduces binary non oriented graphs.

This thesis takes place in this context of determination of effective connectivity, and so in the building of propagation graphs, to detect and quantify the relations between cerebral structures involved in the initiation and the diffusion of the epileptic activities. We propose an evaluation of linear and nonlinear methods, some of them being original. To characterize the proposed algorithms, this evaluation relies on realistic physiological situations. It is completed using simulated autoregressive signals and also real signals recorded on a guinea-pig brain.

This upstream work is a starting point of a more systematic analysis on a larger database of real signals, the actual physiological models used in the present work being constructed to simulate the activity of neuronal populations implied in the propagation of an epileptic seizure. It constitutes a first step in the identification and the localization of modifications in connectivity with a view to diagnosis in terms of the epileptic network organization.

This manuscript is organized as follows: the first chapter allows us to familiarize with the clinical context and recalls some notions of brain anatomy before going on with the problem statement in the clinical context. Even if it is not exhaustive, the chapter 2 is devoted to the most relevant research found in the literature in the field of connectivity as well as to the role played by the

neurocomputational models of epilepsy. In chapter 3, we describe the experimental protocol and detail, on the one hand, the signals we retained in this study, and, on the other hand, the methodology of evaluation to test our approaches. These approaches are developed in chapter 4 and their performance analyzed in chapter 5. Finally, we conclude on this work and suggest future work to this study.

Chapter 1

Clinical context and problem statement

1.1. Human brain

Understanding the internal organization of brain tissue by grasping its logic is a central issue in modern neurobiology. However, the human brain is an extremely complex living structure: its heterogeneity is expressed by the juxtaposition of different territories whose functions are more or less well specified [Boutillier 2007b].

1.1.1. *Location*

The brain is the highest tier in the hierarchy of functional central nervous system and it is especially developed in humans. It is placed in the skull where it rests on the skull base and is covered by a vault.

1.1.2. *External morphology*

From a morphological point of view, the brain looks like a large anterior-posterior oval. It consists of two hemispheres separated by a deep median groove, the inter-hemispheric fissure, and interconnected by bridges of nervous tissue, the inter-hemispheric commissures. The surface of the hemispheres consists of a pallium of very wrinkled grey matter, while at the level of the spinal cord and brain stem, grey matter is central (around the central canal of ependyma). This superficial grey matter is the cerebral cortex or cortex. It has many folds where the deepest ones are called sulci or fissures.

Considering the side of the cerebral hemisphere, in humans, there are three major sulci or fissures on the side of each hemisphere (see Fig. 1.1): the central sulcus or fissure of Rolando, the lateral sulcus or Sylvian fissure, the transverse occipital sulcus or parietal-occipital fissure, a rude sulcus on the external face.

The fissures delimit four lobes, the frontal lobe, the parietal lobe, the temporal lobe, and the occipital lobe. The surface of the lobes is covered by shallow grooves. These furrows delimit large

folds of grey matter called cerebral convolutions (cerebral convolution = gyrus). Thus, the ascending frontal convolution is called pre-central gyrus, and the ascending parietal convolution post-central gyrus. The edges of the lateral sulcus conceal a deep depression - the pit side - which contains a particular lobe called the insula, which displays five small convolutions. The function of this deep lobe appears to be related to conscious sensibilities of visceral origin.

Considering the medial aspect of the hemisphere, it is visible only by cross-section of hemispheric commissures (very specifically by section of the corpus callosum) and opening of the third ventricle.

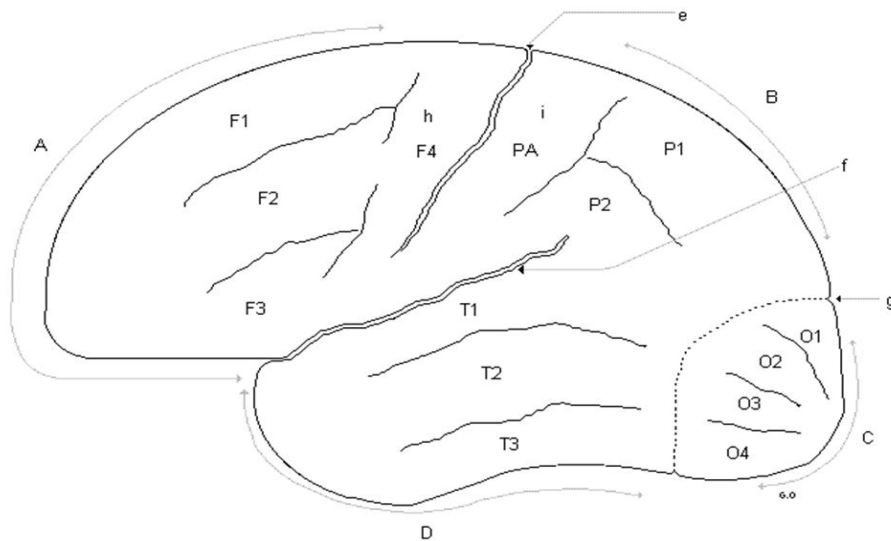


Fig. 1.1 - Lateral surface of the cerebral hemisphere. A: frontal lobe, B: parietal lobe, C: occipital lobe, D: temporal lobe, f: lateral sulcus, e: central sulcus, g: parietal-occipital sulcus, h: pre-central gyrus, i: post-central gyrus [Boutillier 2007a].

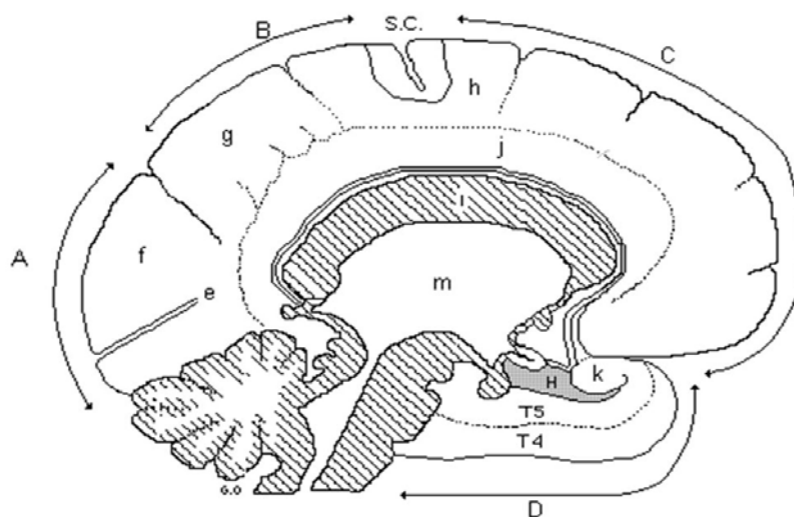


Fig. 1.2 - Medial surface of the cerebral hemisphere. A: occipital lobe, B: parietal lobe, C: frontal lobe, D: temporal lobe, S.C.: central sulcus, H: hippocampus, e: calcarine sulcus, f: cuneus, g: quadrilateral lobule, h: para-central lobule, i: callosum corpus, j: cingulate gyrus, k: uncus, T4: fourth temporal gyrus, T5: fifth temporal gyrus [Boutillier 2007a].

On the medial side (see Fig. 1.2), we distinguish a particular cortical convolution, called the limbic or cingulate gyrus bounded by the cingulate sulcus (calloso-marginal fissure). This convolution is wrapped around the deep part of the hemisphere. The lowest part of the limbic gyrus is wound on itself. It is formed at the bottom by the fifth temporal gyrus, whose end is hook-shaped coil (uncus). It is located against a deeper downturn, called hippocampus. This is a reverse convolution, folded towards the interior of the brain, relief-forming in the cavity of the lateral ventricle. This region of the hippocampus contains the functional structures of memory. Above the limbic gyrus, there is the frontal lobe. On the medial parietal lobe is the quadrilateral lobule and the medial occipital lobe is the cuneus, bounded by the parietal-occipital sulcus (internal perpendicular fissure) and the calcarine sulcus, corresponding to the area of cortical projection of vision.

1.1.3. Internal morphology

The brain contains two types of tissue: the central white matter and the grey matter, the first one being composed of myelin sheath and the second one of neurons and synapses. This grey matter divides up in neocortex and basal nuclei (or basal ganglia) that are situated at the base of the forebrain and strongly connected with the cerebral cortex, thalamus and other brain areas. Finally, in the center of the brain is the ventricular system which is a set of structures containing cerebrospinal fluid in the brain.

1.1.3.1. The cerebral cortex or grey matter

The grey bark is the neocortex. It is around 4 mm thick and covers the entire outer surface of the hemispheres and sinks between the lips of fissures and furrows. The structure of the neocortex is relatively uniform (also called "iso-" and "homotypic" cortex): it consists of nerve cells arranged in six horizontal layers segregated principally by cell types and neuronal connections (Fig. 1.3).

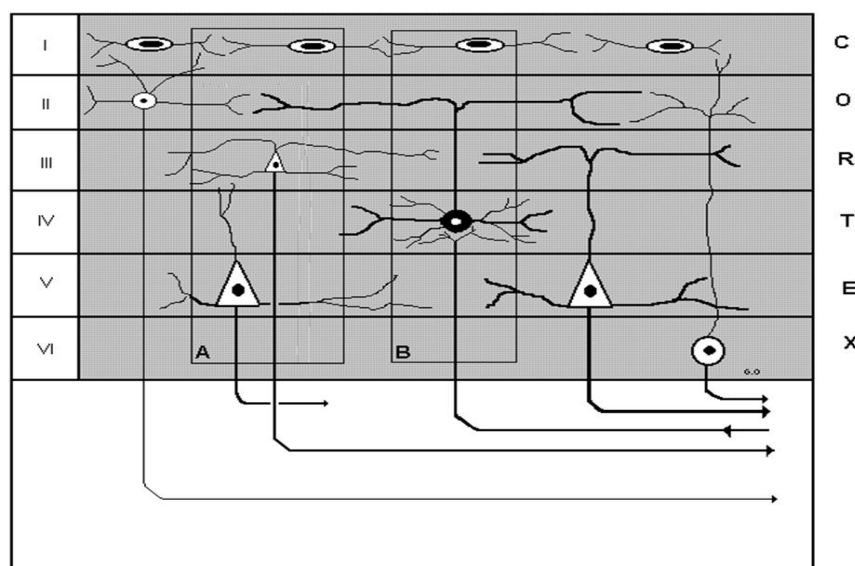


Fig. 1.3 - Architectonics of the cerebral cortex. I: cell surface association, II: cells of intra-hemispheric association, III: small pyramidal cells, IV: projection of sensitive and sensory cells, V: large pyramidal cells of Betz (origin of pyramidal tract), VI: inter-hemispheric cells association (callosal fibers), A and B: functional structure columns [Boutillier 2007a].

The neocortex contains two primary types of neurons, excitatory pyramidal neurons (~80% of neocortical neurons) and inhibitory interneurons (~20%). The first layer is a molecular one composed of scattered neurons, axons and dendrites, the second one contains mainly small pyramidal neurons and numerous stellate neurons. Layer III contains essentially small and medium-size pyramidal neurons. Neurons in layer IV receive all of the synaptic connections from outside the cortex (mostly from thalamus), and this layer distributes all incoming sensory information to the other layers for further processing. Layer V contains large pyramidal neurons and layers V and VI project primarily out of the cortex, e.g. to the thalamus, brainstem, and spinal cord. The neurons of the neocortex are also arranged in vertical structures called neocortical columns, each column responding to a sensory stimulus representing a certain body region of sound or vision. These columns can be seen as the basic functional units of the neocortex.

1.1.3.2. White matter

It is a tissue of the central nervous system and constitutes the internal part of the brain and the superficial part of the spinal cord. It occupies the space between the cortex, basal ganglia and ventricles (Fig. 1.4).

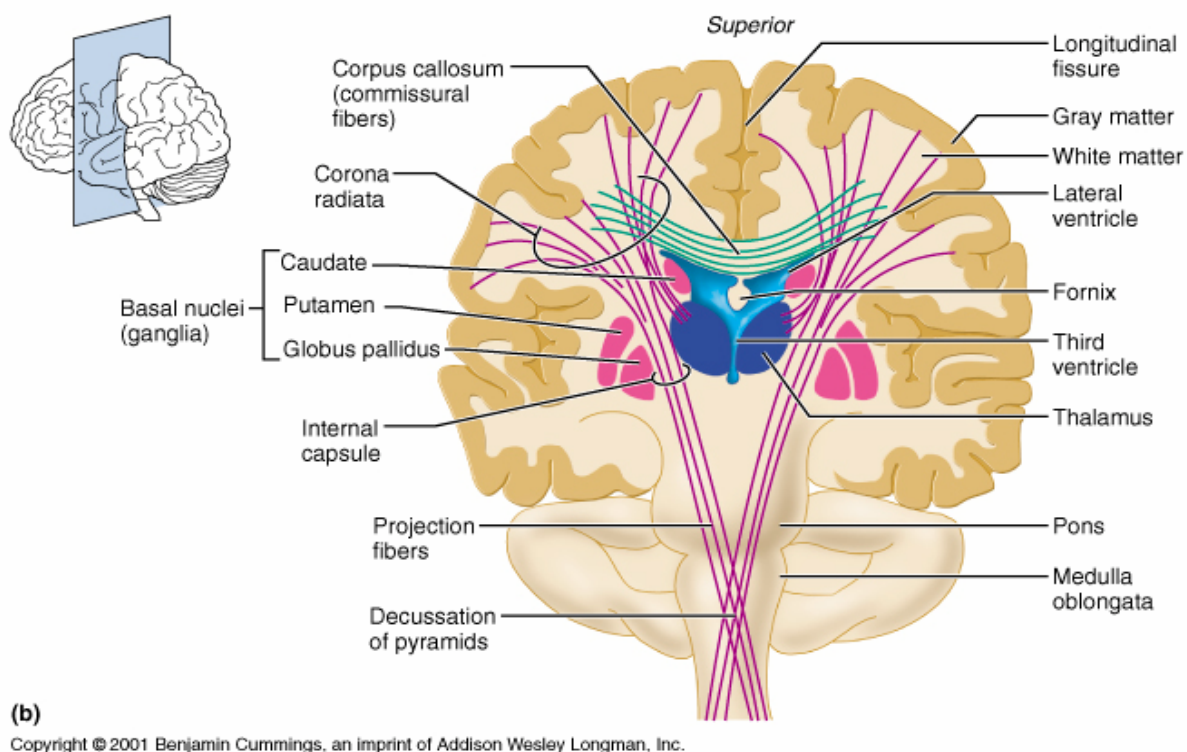


Fig. 1.4 – Brain: matter and nuclei (extracted from [Website1 2001])

The white matter is composed of nervous fibers, myelin-coated axons of neurons. It is a link between different areas of grey matter composed of nuclei of neurons and modulates the distribution of action potentials, acting as a relay and coordinating communication between different brain regions. It transmits signals not only from one region of the cerebrum to another but also between the cerebrum and lower brain centers. There are three different kinds of tracts within the white matter:

projection tracts that carry information between the cerebrum and the rest of the body, commissural tracts that cross from one cerebral hemisphere to the other one, and association tracts that connect different regions within the same hemisphere of the brain. Myelin acts as an insulator, increasing the speed of transmission of all nerve signals.

1.2. Epilepsy

1.2.1. Generalities

1.2.1.1. Prevalence

The prevalence of a disorder is the proportion of a population with that disorder at a given time instant. From many studies around the world, it has been estimated that the prevalence of active epilepsy (*i.e.* continuing seizures needing a treatment) is approximately in the range 4-10 per 1000 of the general population (approximately 1% of the population in France) [Forsgren 2005, Sander 2003]. However, this may be an underestimate as some studies in developing countries (such as Colombia, Ecuador, India, Liberia, Nigeria, Panama, United Republic of Tanzania, Venezuela and so on) suggest a prevalence of more than 10 per 1000 people according to the data provided by World Health Organization (WHO). Thus, it is likely that around 50 million people in the world have epilepsy at any one time. The lifetime prevalence of epilepsy (*i.e.* the number of people presently in the world who have epilepsy now or have had it in the past or will experience it in the future) is approximately 100 million people [Sander 2003].

1.2.1.2. Incidence

The incidence of a disorder is the number of new cases at a given time. Studies in developed countries suggest an approximate annual incidence rate of epilepsy of 40-70 per 100000 of the general population, and higher in infants and elderly people [Forsgren 2005, Macdonald 2000]. However, studies in developing countries suggest that this figure is nearly triple (*i.e.* annual incidence rate of epilepsy of 100-190 per 100000 people) [Mac 2007]. One of the main reasons for the higher incidence of epilepsy in developing countries is the higher risk of experiencing a condition which can lead to permanent brain damage. These conditions include poor sanitation, inadequate health delivery systems, and brain infections and infestations, for instance meningitis, malaria, pre and perinatal complications and malnutrition, ... [Banerjee 2009, Knobler 2004]. In industrialized countries childhood incidence has fallen over the past three decades, which could be a result of adoption of healthier lifestyles by expectant mothers, improved perinatal care, and immunisation programs. On the contrary the rise in incidence in elderly people could be related to improved survival in people with cerebrovascular disease and cerebral degeneration [Everitt 1998, Sander 2003].

1.2.1.3. Mortality

Epilepsy is associated with an increased risk of mortality. Death may be related to factors such as an underlying brain disease (tumor or infection), sudden cardio-respiratory arrest during a seizure, seizures in dangerous circumstances (*e.g.* when swimming), status epilepticus (state of persistent seizure) [Gaitatzis 2004, Sperling 1999, Tomson 2005]....

1.2.2. Historical overview

The word "epilepsy" comes from the Greek words *epi* meaning "upon", and *Leptos* meaning "seizure". It is a common chronic neurological disorder characterized by paroxysmal brain dysfunction due to excessive neuronal discharge, usually associated with some alteration of consciousness, and it affects people in every country throughout the world [Iläe 1993].

In the past, epilepsy was associated with religious experiences and even demonic possession. The main concepts surrounding epilepsy were refined and developed within the magical powers of ancient American Indians during Vedic period (4500-1500BC). In the Ayurvedic detailed literature of Charaka Samhita (400BC), epilepsy was described as a "loss of consciousness". Contrarily to the Ayurvedic medicine of Charaka Samhita, another ancient and detailed interpretation of epilepsy is on a Babylonian textbook (2000BC) in the British Museum in London. It emphasizes the supernatural nature of epilepsy and treatment was, therefore, largely a spiritual matter [W.H.O. 2001]. In Tanzania and in other parts of Africa, epilepsy is associated with possession by evil spirits, witchcraft, or poisoning and is believed by many to be contagious [Jilek-Aall 1999]. In ancient Rome, epilepsy was seen as a curse from the gods. In ancient Greek, it was known as "the sacred disease" because people thought that epileptic seizures were a form of attack by demons, or that the visions experienced by persons with epilepsy were sent by the god or goddess. Hippocrates (400BC), however, believed that epilepsy was not sacred, but was a disorder of the brain. He recommended physical treatments and stated that if the disease became chronic, it was incurable. While both Hippocrates and the Charaka Samhita provided the less spiritualized understanding, the perception that epilepsy was a brain disorder did not begin to take root until the 18th and 19th Centuries AD. The intervening 2000 years were dominated by more supernatural views. Throughout this time, people with epilepsy were viewed with fear, suspicion and misunderstanding and were subjected to enormous social stigma. People with epilepsy were treated as outcasts and imprisoned. Some, however, succeeded and became famous. This is the case of Julius Caesar, Czar Peter the Great of Russia, Pope Pius IX, the writer Fedor Dostoyevsky and the poet Lord Byron.

In the 19th Century, as neurology arose as a new discipline distinct from psychiatry, the concept of epilepsy as a brain disorder became more widely accepted, especially in Europe and the United States of America (USA). This helped to reduce the stigma associated with the disorder. Bromide, introduced in 1857 as the world's first effective anti-epileptic drug, became widely used in Europe and the USA during the second half of the last century and a hospital for the "paralyzed and epileptic" was established in London in 1857. The foundation of our modern understanding of the pathology seen in epilepsy was also laid in the 19th Century with the work of Hughlings Jackson. In 1873, this neurologist proposed that seizures are transient of signs and/or symptoms of abnormal, excessive or hypersynchronous electro-chemical discharges in a group of brain cells and that different parts of the brain can be the site of such discharges [Fisher 2005]. Soon afterwards the electrical excitability of the brain in animals and man was discovered by David Ferrier in London, Gustav Theodor Fritsch and Eduard Hitzig in Germany. Working in Germany during the 1920s, Hans Berger, a psychiatrist, developed the human electroencephalograph (EEG (ElectroEncephaloGraphy) - brainwaves) and its important application was in the field of epilepsy. The EEG revealed the presence of electrical

discharges in the brain. It also showed different patterns of brainwave discharges associated with different seizure types. The EEG also helped to locate the site of seizure discharges and expanded the possibilities of neurosurgical treatments, which became much more widely available from the 1950s.

During the first half of the 20th century, the main drugs for treatment were phenobarbitone (first used in 1912) and phenytoin (first used in 1938). Since the 1960s, there has been an accelerating process of drug discovery, based in part on a much greater understanding of the electrochemical activities of the brain, especially the excitatory and inhibitory neurotransmitters. As we will see in the following, another recent stimulus towards the understanding and treatment of epilepsy in the last few decades has been the development of neuroimaging equipment. Such technology has revealed many of the more subtle brain lesions responsible for epilepsy. Any type of brain lesion (for example, trauma, congenital or developmental infection, degenerative tumor, ...) might lead to epilepsy in some people.

Since stigma is the same in both developed and developing countries, this brain disorder is still hidden and people prefer not to reveal or discuss their condition. Nevertheless, during the last decades greater attention has been paid to quality of life, *i.e.* psychological and social issues, for people with epilepsy, although progress is slow and services are still poor. We must recognize that most of the advances in developed countries are of relevance but not available for the 80% of people with epilepsy who live in developing countries. Of the 50 million people in the world with epilepsy, some 35 million have no access to appropriate treatment. This is either because services are non-existent or because epilepsy is not viewed as a medical problem or a treatable brain disorder. In 1997, the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE) joined forces with the WHO to establish the Global Campaign Against Epilepsy to address these issues. The aim of the ILAE/IBE/WHO Global Campaign Against Epilepsy is to improve prevention, treatment, care and services for patients, the final objective being a favorable environment in which people with epilepsy can live better.

1.2.3. Etiology

As indicated before, epilepsy is one of the oldest conditions known to human beings, and nowadays it is recognized that epilepsy is a neurological disorder characterized by occurrence and recurrence seizures and defined by two or more unprovoked seizures [Blume 2001, Ilae 1993, Levisohn 2007]. These seizures are the result of sudden, usually brief, electro-chemical discharges in the brain, and their character depends on the location and function of the site of the discharges [Fisher 2005]. The clinical manifestations of seizures therefore vary and depend on where in the brain the disturbance first starts and how far it spreads. Transient symptoms can occur, such as loss of awareness or consciousness and disturbances of movement, sensation (including vision, hearing and taste), mood or mental function.

Seizures may vary from the briefest lapses of attention or muscle jerks to severe and prolonged convulsions. They may also vary in frequency, from less than one a year to several per day. Seizures are classified according to where they arise in the brain [E.F.A. 2012, Engel 2001, 2006, Olson 2008, T.N.S.E. 2009], for instance:

- Partial or focal seizures (localization-related seizures): These seizures are further divided on the extent to which consciousness is affected. If it is unaffected, then it is a simple partial seizure; otherwise it is a complex partial psychomotor seizure [Mckee 2010]. These seizures arise from an electrical discharge of one or more localized areas of the brain, but then they may spread to the whole brain arousing a generalized seizure.
- Generalized seizures (distributed seizures): The electrical discharge which stirs up these seizures affects both sides of the brain and may lead to loss of consciousness and/or muscle contractions or stiffness. Generally, they include absence, myoclonic, clonic, tonic, tonic-clonic, and atonic seizures [Daly 1968, Tramonte 2008].
- Status epilepticus (unknown localization): This is a state in which a person has frequent seizures without recovery of consciousness between each episode. It is also a dangerous state and if not treated may result in brain damage or death.

Note that children may exhibit behaviors that are easily mistaken for epileptic seizures but are not caused by epilepsy. These include inattentive staring, benign shudders (among children younger than age 2, usually when they are tired or excited), or self-gratification behaviors (nodding, head banging). Sometimes, these behaviors can be distinguished from epilepsy because the episodes never occur during sleep, for example [Olson 2008].

Just as there are many types of seizures, there are many types of epilepsy syndromes. Epilepsy classification includes more information about the patient and the episodes than seizure type alone, such as clinical features (e.g. behavior during the seizure) and expected causes [E.F.A. 2012].

There are more than forty different types of epilepsy, including benign Rolandic epilepsy, frontal lobe epilepsy, infantile spasms, juvenile myoclonic epilepsy, Lennox-Gastaut syndrome, Landau-Kleffner syndrome, mitochondrial disorders, progressive myoclonic epilepsy, reflex epilepsy, Rasmussen's syndrome, temporal lobe epilepsy, limbic epilepsy, Rett syndrome, status epilepticus, Jacksonian seizure disorder, Lafora disease [T.N.S.E. 2009].

Until now, it is unclear why particular seizures occur at a particular age or time and not at other ages or times. Provocative factors, however, are recognized in some patients. For example, certain flashing lights (discos, television, video games), over-breathing, over-hydration, loss of sleep, and/or emotional and physical stress, may stimulate seizures. Although these are not causes of epilepsy, they may influence the timing and frequency of seizures and so far we know that there are different causes of epilepsy that are common in certain age groups [Fauci 2008]:

- In the neonatal period and early infancy the most common causes include hypoxic-ischemic encephalopathy, Central Nervous System (CNS) infections, trauma, congenital CNS abnormalities, and metabolic disorders
- In late infancy and early childhood, febrile seizures are fairly common and may be caused by CNS infections and trauma

- Generally, during adolescence and adulthood, the causes are secondary to CNS lesion. Further, idiopathic epilepsy is less common. Other causes associated with these age groups are stress, trauma, CNS infections, brain tumors, and illicit drug use and alcohol withdrawal
- In older adults, cerebrovascular disease is a very common cause. Other causes are CNS tumors, head trauma, and other degenerative diseases that are common in the older age group, such as dementia.

Even if an underlying brain disease can cause epilepsy, all people with the same brain disease will not necessarily have epilepsy. In view of the fact that only a proportion of people who have a brain disease experience seizures as a symptom of that disease, it is suspected that people who do have such symptomatic seizures are more vulnerable due to biochemical/neurotransmitter reasons. Epileptic syndromes differ from the age of onset, the type of seizure, the presence or absence of detectable brain disease and genetic background. However, medical science is only at an early stage in understanding these different types.

1.3. Towards solutions

Since epilepsy has no age, racial, social, sexual or geographical boundaries [Hirtz 2007] and can begin at any age, anyone can be affected by seizures. In fact, up to 5% of the world's population may have at least a single seizure at some time in their lives, but a diagnosis of epilepsy is reserved for those who have recurring seizures, at least two unprovoked ones.

1.3.1. Scientific and medical advances

Research on epilepsy is categorized in two groups: fundamental and clinical research. In today's society, health care is a privilege and the majority of the technology devoted to this research and the benefits which accrue from it are only available in developed countries.

1.3.1.1. Fundamental research

This focuses on the fundamental mechanisms which underlie the development of epilepsy, the cause of spontaneous seizures, their different manifestations, their timing and duration, and the consequences of repeated seizures on brain function. Understanding the cellular and brain processes responsible for individual seizure types and epileptic disorders will lead to new approaches to prevention, treatment and care. Recent work has traced specific types of seizures to distinct disturbances in neuronal connections in the brain, and the chemical transmission of information between neurons. Neuropharmacologists have subsequently designed compounds which selectively interfere with these abnormal brain functions, leading to the development of new anti-epileptic drugs which are able to treat specific types of epilepsy with less impairment of normal brain function. For instance, they are less sedative and have fewer cognitive side-effects. In the past few years, research has made advances on the genetic basis of some epileptic syndromes, mainly in childhood and adolescence, with the identification of specific chromosomal linkages which increase the probability that an epileptic disorder will appear, usually in association with other acquired or environmental

factors. Identification of some of the genes responsible for a predisposition to epilepsy may reveal the basic neurochemical or physiological defects which need to be prevented or corrected. This, in turn, may help scientists to develop new anti-epileptic treatments. We can expect that, in the next decade, research on molecular genetics of human epilepsy will result in an entirely new classification of epileptic disorders and a better understanding of the fundamental causes of many forms of epilepsy.

1.3.1.2. Clinical research

This research is based on new diagnostic and therapeutic technologies and includes understanding differences in the various types and causes of epilepsy by neuroscientists. This understanding has been improved by the help of researchers in the fields of psychology and sociology enabling the formulation of rehabilitation programs. It is only in the last decade that clinical neuroscientists have been able to look directly at the structure and function of the living human brain. This has been done thanks to Magnetic Resonance Imaging (MRI) that has enabled to visualize structural brain abnormalities responsible for epileptic seizures, Positron Emission Tomography (PET) and Single Photon Emission Computed Tomography (SPECT) that help in pinpointing an epileptic region by looking at localized dysfunction in brain blood flow, metabolism and chemical processes during and between seizures, ElectroEncephaloGraphy (EEG) and MagnetoEncephaloGraphy (MEG) that give information on the sites of origin of epileptic discharges, Magnetic Resonance Spectroscopy (MRS), that is a non invasive technique to identify areas of brain damage as well as disturbances in brain metabolism and neurotransmitter function. Most of these techniques are used in epilepsy centers in developed countries not only for research but also for evaluation of people who may benefit from brain surgery as treatment for intractable, drug resistant epilepsy.

Therapy for epilepsy includes two main aspects: pharmacotherapy and surgical therapy. Pharmacological and surgical therapies are based on correcting or eliminating specific epileptic disturbances. This not only provides the greatest opportunity for people to be relieved of disabling seizures and the disturbing side-effects of treatment, but it also avoids the long-term psychosocial consequences of living with an uncontrolled epileptic disorder.

1.3.1.2.1. Pharmacotherapy

AntiEpileptic Drugs (AEDs) are the most common treatment of epilepsy. Often, anticonvulsant medication treatments are lifelong and have major effects on quality of life. At present, in the USA, there are twenty medications approved by the Food and Drug Administration (FDA) for the use of treatment of epileptic seizures. Most of common antiepileptic drugs came into the market between 1960 and 1990. For approximately two decades, no new major anti-epileptic medication was introduced but several drugs are currently in the clinical testing stages. Some drugs are commonly used to abort an active seizure or interrupt a seizure flurry and some others are used only in the treatment of refractory status epilepticus.

Mechanisms, effectiveness for particular epilepsy syndromes, and side-effects differ among the individual anticonvulsant medications [Marson 1996]. It is difficult to find a treatment that is totally effective and about 20% of patients with epilepsy continue to have breakthrough epileptic seizures

despite best anticonvulsant treatment [Cascino 1994, Engel 1996]. Moreover, 88% of patients with epilepsy, in a European survey, reported at least one anticonvulsant related side-effect [Baker 1997, Baker 2000]. Most side-effects such as mood changes or sleepiness depend on the dose of the drug and can often be avoided or minimized by the use of the smallest effective amount. Now, the objective goal for individual patients is no further seizure and minimal side-effects, and the job of the physician is to aid the patient to find the best balance between the two states. Most patients can achieve this balance with monotherapy, *i.e.* the use of a single anticonvulsant medication while some others require polypharmacy, *i.e.* the use of two or more anticonvulsants.

Generally, we can expect that, in 50% of cases, the seizures can be controlled using a first single drug. This ratio can reach 65% when using a second or a third drug or even 70% when using two drugs simultaneously. After two to five years of successful treatment, drugs can be withdrawn in about 70% of children and 60% of adults without relapses [Brodie 2005, French 2007, Kwan 2000]. Other treatments, in addition to (or instead of) anticonvulsant medications may be considered by those people with continuing seizures. Attention is now being directed at the most appropriate choice of drug for specific epilepsy syndromes. If pharmacotherapy has no beneficial effect, the epilepsy is said to be medically refractory and a surgery is recommended. Today the reason why some people develop chronic drug-resistant epilepsy is also a field of investigation.

1.3.1.2.2. Surgical therapy

Since the last twenty years, it has been recognized that specific epileptic syndromes responding poorly to drug treatment have an excellent chance of successful treatment with surgical intervention. Epilepsy surgery is an option for patients whose seizures remain resistant to treatment with anticonvulsant medications who also have symptomatic localization-related epilepsy, a focal abnormality that can be located and therefore removed [Duncan 2006]. The goal for these procedures is to totally control the epileptic seizures, although AEDs may still be required [Berg 2007, Birbeck 2002, Loring 2004].

In epilepsy surgery, the goal is to locate the epileptic focus and to determine if resective surgery will affect normal brain function. The evaluation for this surgery includes neurological examination, EEG recording, long-term video-EEG monitoring, neuropsychological evaluation, and neuroimaging such as MRI, SPECT, PET. Some epilepsy centers also use intracarotid sodium amobarbital test (Wada test), functional MRI for the Blood-Oxygen-Level-Dependent (BOLD) contrast or MEG as supplementary tests [Adcock 2003, Benke 2006, Liégeois 2004]. Moreover, invasive intracerebral recordings can be needed and employed with long-term video-EEG monitoring when noninvasive testing appears inadequate to identify the epileptic focus or distinguish the surgical target from normal brain tissue and function. Techniques such as ElectroCorticoGraphy (ECoG) can also be used in this context. Generally, the most common surgery consists in the resection of lesions like tumors or malformations, which, in the process of treating the underlying lesion, often result in control of epileptic seizures caused by these lesions. The most common form of intractable epilepsy in adults is temporal lobe epilepsy with hippocampal sclerosis, and the most common type of epilepsy surgery is the anterior temporal lobectomy, or the removal of the front portion of the temporal lobe including the

amygdala and hippocampus. Even if this surgery is effective, durable, and results in decreased health care costs, patients can decide not to undergo such a surgery due to fear in having a brain operation. Another type of surgery refers to palliative surgery to reduce the frequency or severity of seizures. It is recommended to patients when there is no other solution to control the seizures. Examples are callosotomy or commissurotomy to prevent seizures from generalizing in the entire brain, which results in a loss of consciousness. When the epileptic focus is located near important functional areas of the cortex, the surgeon can prescribe multiple subpial transections to decrease the spread of seizures across the cortex. This solution consists in carrying out cuts to interrupt fibers that connect neighboring parts of the brain. For the most catastrophic epilepsies, people must undergo hemispherectomy which is the most invasive surgical operation to remove one cerebral hemisphere. It is reserved for people whose seizures have not responded to medications or other less invasive surgeries. After the operation, most patients suffer from hemiplegia on the side of the body opposite the removed or disabled portion.

1.3.2. Problem statement

As presented above, in around 30% of the cases, epilepsies remain drug-resistant. They are often partial or focal, with an origin located in a relatively circumscribed area of the brain (temporal, frontal, parietal, ...). For these partial epilepsies, a surgical operation can be considered. The difficulty that arises is then to determine, for a given patient, which areas of the brain should be resected such that seizures are suppressed or strongly attenuated, under the constraint that post-surgical deficits (sensitive, driving or cognitive) induced by surgery are limited. Therefore, the problem is to define, from observations, the epileptogenic zone that is responsible for seizures. Indeed, the organization of this epileptogenic zone often corresponds to that of a network of neuronal ensembles distributed in distant structures and responsible for the initiation and the propagation of epileptic activities. Due to this distributed feature, the localization and the characterization of epileptogenic networks is a difficult problem. Both are however crucial because they can lead to finer delimitation of the cerebral volume to be removed (resection) in order to suppress epileptic activity.

In this context, one understands the challenge that represents the analysis of intracerebral electroencephalographic signals to (i) identify the cerebral structures implied in the different phases of a seizure (*i.e.* identify the nodes of a graph), (ii) quantify the information carried by the different signals (attributes for each node), and (iii) characterize the temporal inter-structures relations thanks to the joint analysis of these signals (statistical links and detection of causal influences). The present study aims at giving some answers to the latter point, *i.e.* quantifying information flow between structures in detecting/analyzing relations inside an ensemble of signals recorded on multiple channels, to better understand the organization of the seizure in terms of origin and propagation. The non stationary character of this organization is taken into account under the constraint of short analysis windows (only few seconds) to cope with coupling evolution.

To this end, our study is devoted to the analysis of signals that are characteristic of epileptic activity, either simulated or *ex vivo* activity. Developments are first applied to simple simulated models generated by Vectorial AutoRegressive (VAR) models, before testing physiology-based models and

finally real epileptic signals. For the last class of signals, they consist of signals recorded during epileptic seizures induced in a guinea-pig, since the recorded activity shows some common characteristics with that recorded in humans at the beginning of the seizure. The animal model comes from data recorded in the Unit of Experimental Neurophysiology and Epileptology, Fondazione Istituto Neurologico, Milan, by Pr. M. de Curtis. The investigated technique consists in inducing epileptic seizures by perfusing bicuculline in isolated brains (maintained *in vitro* during a dozen of hours) of guinea-pigs, and recording the resulting electrical activity through the use of 16-channel silicon probes inserted in the tissue as perpendicular as possible to the entorhinal cortex, this cortex playing some crucial role in temporal seizures initiation as confirmed by studies in humans.

At present, it seems that five phases can be considered in the seizure, interictal spikes, pre-ictal spikes, ictal onset corresponding to a narrowband tonic activity (FOA: Fast Onset Activity), ictal after-discharges, and post-ictal spikes. Now, we would like to know the cerebral structures implied in the seizures according to the different phases and more particularly during the FOA.

To answer this question, beyond the functional connectivity relative to the statistical coupling between signals recorded in given areas, we have to infer the effective connectivity, *i.e.* elaborate graphs of information flows between concerned populations, from these signals. To this end, various techniques have been already proposed in the literature. More often, we observe possible insufficiency of linear techniques versus nonlinear ones. Some of them that are non parametric are based on marginal and joint entropies, such as mutual information if we are concerned with functional connectivity, or on transfer entropy if we are concerned with effective connectivity. Right now, we sense the difficulty (i) to get a "good" connectivity index in terms of asymptotic performance but also in terms of robust estimation on short time samples, (ii) to validate performance both in simulation, since we work only on "models" (even if they are realistic), and in real situations, since no "ground truth" is available.

In this document, we try to give some elements and we propose to investigate linear and nonlinear techniques, some of them being original, and to apply them to real signals, or at least to simulated signals using an argued physiology-based model. Partial responses will be given to the choice of the "best" index to discover (i) which structures are implied in rapid activity, (ii) which functional/effective connectivity exists between observed areas, and (iii) which structures "drive" other ones at the seizure onset. The long-term objective of this research is to help to understand the mechanisms of the seizure in order to avoid it, abort it, or, at least, to stop it or interrupt it.

Chapter 2

Connectivity and models

In my work, I had to investigate methods aimed at analyzing causal connectivity between epileptic signals in the human brain. Therefore, to give a scope of the application context, and after the brief description of brain anatomy introduced in chapter 1, two main aspects are discussed in this chapter. The first one concerns brain connectivity, whose fundamental concepts and definitions (or pseudo-definitions) are presented next, as well as the instrumental and algorithmic methods actually available to investigate experimentally this paradigm. The second one focuses on the main classes of computational models introduced to investigate epileptic processes and which must be considered not only to obtain physiologically plausible simulated signals in order to test algorithms as done in this thesis but also to help the interested reader to understand model-based inference methods proposed elsewhere in the literature even if they have not been investigated here.

2.1. Brain connectivity

There has been a growing interest in studying both normal and pathological brain functions by identifying variations in activation and interactions between brain areas: understanding and modeling brain functions is based not only on the correct identification of the active brain regions, but also on the functional interactions among the neural assemblies distributed across different brain regions. The aforementioned concepts are addressed in theoretical neuroscience, as the functional segregation (activation of specialized brain regions/neural assemblies) and integration (coordinated activation of very large numbers of neural assemblies distributed across different cortical areas that constitute large-scale distributed systems of the cerebral cortex) principles [Friston 2009b].

Integration of cerebral areas can be measured by assessing brain connectivity. Brain connectivity refers not only to a pattern of anatomical links ("anatomical or structural connectivity"), but also to statistical dependencies ("functional connectivity") or more specifically to causal interactions ("effective connectivity") between distinct units within a nervous system. The units may correspond to individual neurons, neuronal populations, or anatomically segregated brain regions. From a synthetic

point of view, connectivity pattern is conditioned, but not completely determined, by structural links such as synapses or fiber pathways. Statistical or causal relationships measured as cross-correlations, coherence, or information flow must be taken into consideration. Indeed, different classes of dynamical neural processes can live on a same structural substrate, depending on the modification of some parameters (clearly structural modification generally implies dynamical modifications). Neural activity, and by extension neural codes, are constrained by connectivity. Brain connectivity is thus crucial to elucidate how neurons and neural networks process information [Edeline 1999, Sakkalis 2011, Sporns 2004, Sporns 2010, Treserras 2008].

In this section, according to the three aforementioned concepts of connectivity, the different and interrelated aspects of brain organization are first discussed in section 2.1.1 before going on in section 2.1.2 with the observational and computational methods used to explore these connectivities and discussing their relationship in section 2.1.3.

2.1.1. Generalities

2.1.1.1. Modes of brain connectivity

When applied to the brain, the term "connectivity" refers to several different and interrelated aspects of brain organization [Horwitz 2003]. A fundamental distinction in connectivity comes from the notions of structural connectivity, functional connectivity and effective connectivity [Friston 1994, Sporns 2004, Sporns 2005, Sporns 2007].

Anatomical connectivity is viewed as a network of physical or structural (synaptic) connections linking sets of neurons or neuronal elements, as well as their associated structural biophysical attributes encapsulated in parameters such as synaptic strength or effectiveness. At shorter time scales, *i.e.* from seconds to minutes, the physical pattern of anatomical connections can be relatively stable. But at long time scales from hours to days, structural connectivity patterns are more subject to significant morphological change and plasticity.

Functional connectivity generally refers to a statistical concept. Such connectivity captures deviations from statistical independence between events and/or activities concerning distributed and often spatially remote neuronal units. Statistical dependence may be in many cases estimated by measuring well-known correlation, covariance or spectral coherence. More often, this functional connectivity is estimated between all pairs of elements of a system, regardless of the direct or indirect connections (or structural links) between these elements. Unlike structural connectivity, functional connectivity can be highly time-dependent. Statistical patterns of links among neuronal elements fluctuate on multiple time scales, some as short as tens or hundreds of milliseconds. It should be noted that "pure" functional connectivity does not generally make any explicit reference to specific directional effects or to an underlying structural model.

At a third stage, effective connectivity may be considered as a lecture of functional connectivity highlighted by the structural connectivity concept, as it describes networks of directional effects of one neural element over another. It takes account of the fact that neural processes do not violate the causality principle introduced in physics. In principle, causal effects can be inferred through "invasive"

systematic perturbations of the system if possible as in cognitive experiments, or through electromagnetic stimulations of particular brain areas in epileptic patients, but also, since causes must precede time effects, through analysis of time precedence in time series, as we will see later. To detect effective connectivity, some techniques require the specification of a model which encodes causal links and which includes structural parameters, whereas other ones are (more or less) "model-free", for example those involving the application of time series causality measures such as Granger causality or transfer entropy (see chapter 4).

2.1.1.2. Graph theory formalism and techniques for brain connectivity

Brain connectivity may be studied and analyzed using a broad range of network analysis approaches [Rubinov 2010]. In this scope, different ways to represent brain connectivity patterns, such as graphs or matrices like those displayed in Fig. 2.1, appear interesting as they apply to structural, functional and effective brain connectivity at all levels.

Formally, a graph is a collection of vertices and edges. In our context, vertices (or "nodes") correspond to neurons in small-scale connectivity or to brain regions in large-scale connectivity. They can be linked through direct connections (edges), or indirectly via paths composed of multiple edges. An edge can be oriented or not. The graph is called "oriented graph" when each existing edge is oriented. An edge can also be valued by assigning it a value (a real number) which weights the link. Different attributes can be easily defined on a graph. For example, the distance between two vertices is defined as the length of the shortest path between these vertices, and the global average of all distances is called the characteristic path length.

Clearly, according to the different types of connectivity, and given a set of vertices corresponding to distinct neural elements, more than one graph may be designed. A possibility is to consider a purely structural graph with edges (or "links") corresponding to synapses or pathways in small-scale connectivity, or to fibers bundles in a more large scale. Neural dynamical interactions are described with graphs representing functional and effective connections. Since a significant value of functional efficacy can be measured between vertices not directly linked (minimal path length including more than one existing structural edge), a functional connectivity graph generally includes more edges than the corresponding structural graph, when the latter is known. So, when two neural sites are linked by more than one path, some of them may not correspond to a real structural path. Generally, the importance of indirect interactions depends on the path length defined on the structural graph. Concerning effective connectivity, the links of the graph are oriented. Note that, without any exhaustive knowledge of anatomical (structural) links, it can be impossible to decide, from functional or effective connectivity measures encoded in a graph, whether a given edge corresponds (or not) to a real set of axonal fibers. It is possible to detect a functional edge without structural edge, and, conversely, to detect a structural link without significant statistical link between two corresponding nodes.

Graphs may be described by a connection matrix (named also adjacency matrix) with elements that represent the presence or absence of a directed edge between pairs of vertices. For each type of connectivity, a matrix can be hence recovered from the corresponding graph. An example is illustrated

in Fig. 2.1. For anatomical connectivity, we are concerned with sparse matrix and (partially) directed graph since we can generally consider either the presence or the absence of an element (e.g. the presence or not of a fiber or a bundle of fibers). In their pure topologic form, the matrix entries are binary elements indicating the presence or absence of a connection, but edges may also be weighted, where the weights represent connection densities or efficacies. For functional connectivity, the corresponding matrix is symmetric, with each element encoding statistical dependence or proximity between two system components (neurons, recording sites, ...). Such matrices may be thresholded to produce binary undirected graphs, a high threshold resulting in high sparsity. Finally, for effective brain connectivity, unlike functional connectivity, it yields a full non symmetric matrix due to possible unidirectional paths or to bidirectional paths differently balanced. Applying a threshold to such matrices yields also binary directed graphs.

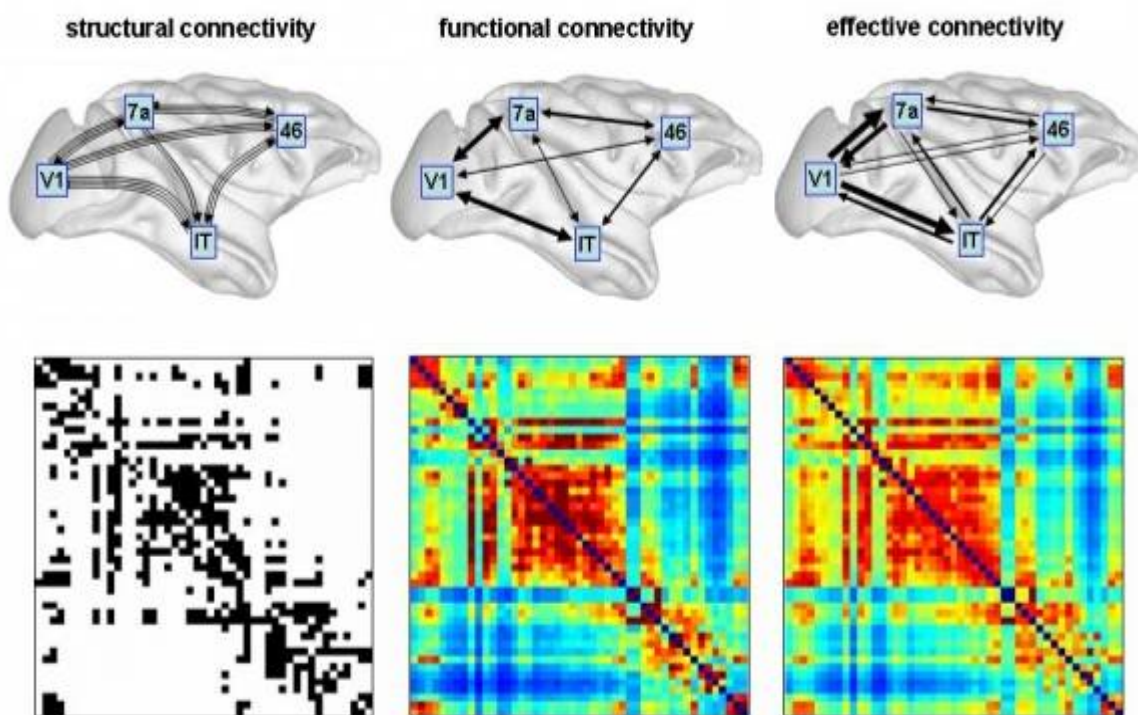


Fig. 2.1 – Modes of brain connectivity. Sketches at the top illustrate structural connectivity (fiber pathways), functional connectivity (correlations), and effective connectivity (information flow) among four brain regions in macaque cortex. Matrices at the bottom show binary structural connections (left), symmetric mutual information (middle) and non-symmetric transfer entropy (right). Data was obtained from a large-scale simulation of cortical dynamics (see [Honey 2007]).

Graphs of brain networks can then be quantitatively described, for example through vertex degree (number of links connected to a node) and strength (links being associated with connection weights), assortativity (degree of correlation between two nodes), subgraphs (motifs underlying the notion of local patterns of connectivity, for instance feedforward loops, feedback loops and bidirectional loops between two nodes), clustering coefficients (measure of degree to which nodes in a graph tend to cluster together), path lengths (distances), or vertex and edge centrality (measuring how influential is a vertex or an edge in a graph) (see Fig. 2.2), among many other graph theory measures [Brandes 2005].

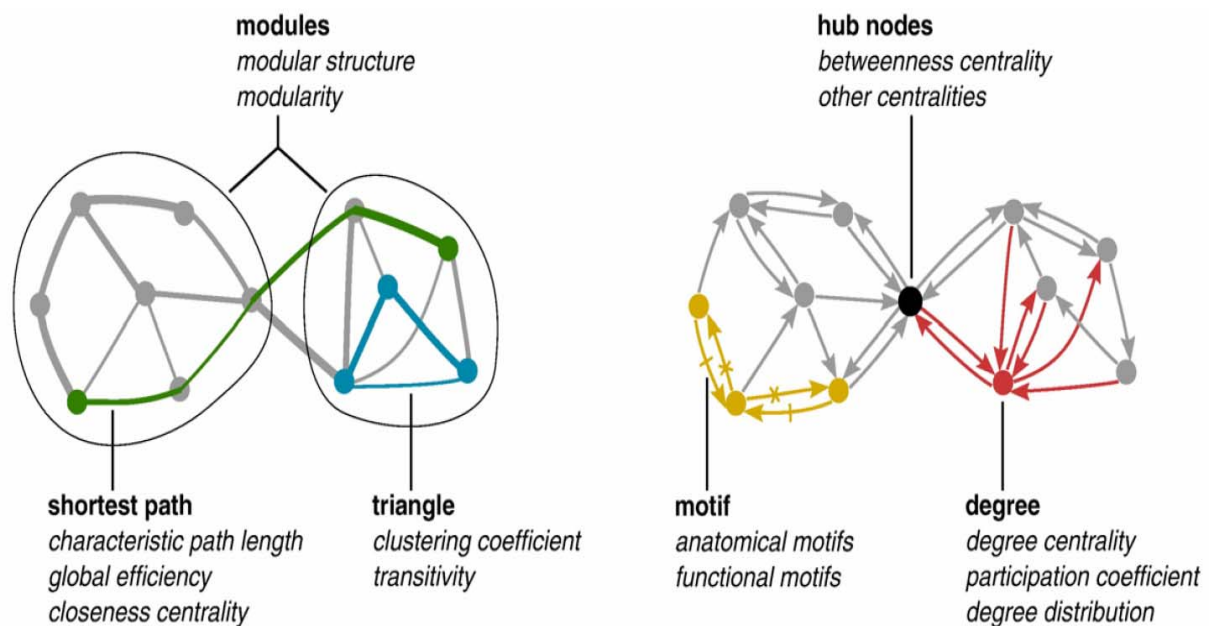


Fig. 2.2 – An illustration of some key complex network measures (in *italics*) for graphs of brain networks. These measures are typically based on basic properties of network connectivity (in **bold type**). Thus, measures of integration are based on shortest path lengths (green), while measures of segregation are often based on triangle counts (blue) but also include more sophisticated decomposition into modules (ovals). Measures of centrality may be based on node degree (red) or on the length and number of shortest paths between nodes. Hub nodes (black) often lie on a high number of shortest paths and consequently often have high betweenness centrality. Patterns of local connectivity are quantified by network motifs (yellow). An example three-node and four-link anatomical motif contains six possible functional motifs, of which two are shown—one motif containing dashed links, and one motif containing crossed links [Rubinov 2010].

To be operational these measures must be considered statistically, requiring the design of appropriate null hypotheses (referring models), involving the choice of suitable random graph models. Such models are not uniquely defined, as statistical comparisons may be carried out relative to a number of different random models that statistically preserve various subsets of structural features. For example, some models involve edge randomization techniques conditioned by given vertex degrees.

Beyond the analysis of the network's topological features, other analysis approaches focus on the three-dimensional structure of brain networks. These approaches include morphometric methods, for example those for measuring wiring length or volume [Wen 2005].

2.1.2. Observational and computational methods

2.1.2.1. Structural connectivity

A difficulty encountered to define structural connectivity is that, at the microscopic scale of neurons, destruction and creation of some synaptic connections can occur dynamically and be largely dependent on the executed function [Ooyen 2001]. Complexity of nervous systems relates to their intricate morphology, and to the rich interconnectivity of the neuronal processing elements. Neural connectivity patterns influence strongly the functional properties of neurons and neuronal systems.

Brain connectivity can be described at several levels of scale. These levels include synaptic connections that link individual neurons at the microscale, networks connecting neuronal populations at the mesoscale, as well as brain regions linked by fiber pathways at the macroscale. At the microscale studies have described many basic components and micronetworks in the cerebral cortex. At the mesoscale, neurons are organized into networks of columns and minicolumns. At the macroscale, distinct brain regions including many neuronal populations are interconnected and form large-scale connectivity patterns.

In the past, first investigations of structural connectivity in the human brain were limited to the use of post-mortem dissection techniques. These techniques only allowed to describe the repartition and the orientation of the most important fiber pathways on a macroscopic level [Klingler 1956, Vieussens 1684]. Afterwards, some myelin staining techniques have allowed macroscopic examinations of white matter bundles [Flechsig 1920]. But these techniques barely allow tracking the three-dimensional course of a particular fiber tract. Therefore, until the end of the 20th century, knowledge on the human brain structural connectivity was mostly limited to the course of major fiber bundles (for example the cortico-spinal tract or the arcuate fasciculus).

More recently and thanks to the progress in diffusion Magnetic Resonance Imaging (dMRI) and in tractography techniques, Diffusion Tensor Imaging (DTI) have been used to extract information concerning the white matter fiber tracts [Basser 1994, Johansen-Berg 2006, Johansen-Berg 2009, Le Bihan 1986]. Attempts have been made to map at the scale of brain regions the human connectome (a connectome is an exhaustive description of the structural connections in a given nervous system) [Sporns 2005]. For example, cross-correlations in cortical thickness across individual brains allowed to indicate the presence of cortico-cortical pathways [He 1998] and a sophisticated technique, namely the Diffusion Spectrum Imaging (DSI), has been used [Hagmann 2007] to detect crossing fiber bundles. Quite recently, Paus *et al.* [Paus 1997] proposed to combine TMS (Transcranial Magnetic Stimulation) and PET to map neural connections in alive human brain. TMS allows to stimulate selectively a given cortical area and to measure simultaneously brain activity modifications, detected by CBF (Cerebral Blood Flow) measured with PET [Ferreri 2011, Hallett 2000].

Globally, all these studies tend to reveal highly clustered large-scale cortical network, where pathways mostly exist between areas that are spatially close. This leads to the conclusion of a small-world connectivity of the cortical network.

2.1.2.2. Functional and effective connectivity

2.1.2.2.1. Overview on instrumental modalities

Functional connectivity is defined as the temporal correlation (in terms of statistically significant dependence between distant brain regions) among activities of different neural assemblies [Fingelkurts 2005]. Many neurophysiological signals can be assessed with functional connectivity techniques, including signals derived from single unit and LFP (Local Field Potential) recordings, EEG, MEG, PET and fMRI. Marginal correlations between two neural assemblies are not the only indicators, partial (conditional) measures being also useful to explain indirect links.

Most of studies in mapping functional networks are based on fMRI and EEG observations [Achard 2006, Achard 2007, Achard 2008, Bassett 2006a, Bassett 2006b, Meunier 2009, Stam 2002a, Stam 2002b], these two approaches being considered as complementary since the first one has good spatial resolution and the second one good temporal resolution. Achard *et al.* [Achard 2006] computed cross correlations from human fMRI data and found that functional brain networks exhibit small-world attributes possibly reflecting the underlying anatomical connections. Concerning EEG, functional connectivity between pairs of electrodes has been estimated using different methods. Stam used generalized synchronization (GS) [Stam 2002a] to extract functional networks from EEG and MEG data sets. In [Bassett 2006a], wavelet correlation allowed to estimate frequency-dependent functional connectivity between MEG sensors. This work showed an invariance of the small-world characteristic of the functional connectivity through different frequency bands.

Effective connectivity has been extensively investigated with multi-modal functional neuroimaging studies [Breakspear 2004, Brovelli 2004, David 2008, Seth 2005, Seth 2007] using various techniques. For example, in [McIntosh 1994], the analysis of brain imaging data focused on the examination of the covariances of activity among neural regions during different behaviors. The technique refers to structural equation modeling, where connections between brain areas are based on *a priori* known neuroanatomy, and the interregional covariances of activity are used to calculate path coefficients representing the magnitude of the influence of each directional path. This structural model-based approach has allowed to identify significant differences in effective connectivity between a given set of brain regions when measured during different cognitive tasks. This illustrated the dependence of these patterns on time and on task.

LFP recordings have been analyzed to measure directional influences (Granger causality) among recording sites and have provided information about directed interactions in the course of behavioral and cognitive tasks [Brovelli 2004].

Transcranial Magnetic Stimulation (TMS) has also been associated to functional neuroimaging [Ferreri 2011, Hallett 2000, Paus 1997]. TMS can hence be used to test how cortex stimulation modifies the activity and the connectivity in the disturbed neuronal circuits. PET, fMRI and EEG have been combined to assess how TMS affects neural processing locally and also in remote interconnected brain regions [Bestmann 2008, Ruff 2009].

Besides studies in healthy volunteers, the combined TMS/PET technique has potential in clinical studies of various neurological and psychiatric disorders for which it is crucial to evaluate the state of functional connectivity independently of the sensory, motor, and cognitive abilities of the patient [Ferreri 2011, Hallett 2000, Paus 1997]. The combination of TMS with acquisition modalities such as EEG [Ferreri 2011] is presented also as a possibility to better characterize the effective connectivity in neuronal cortical networks and also the mechanisms involved in regulating the balance between inhibition and facilitation in the cortex and the corticospinal pathway. TMS with high-density EEG has also allowed to understand how the activation of the premotor area is transmitted to the rest of the brain. Thus, the fading of consciousness during some sleep stages could be related to a breakdown in cortical effective connectivity [Massimini 2005].

2.1.2.2.2. Data processing in functional connectivity

In the following, the signal processing techniques used to measure functional connectivity are divided into two main groups: linear and nonlinear techniques.

2.1.2.2.2.1. Linear techniques

In the 1950s, linear brain connectivity began to be measured using cross-correlation of pairs of EEG signals [Brazier 1952] and partial correlation in trivariate signals [Blalock 1961]. Higher correlations indicate stronger functional relationships between the related brain regions. In order to measure linear connectivity in the frequency domain, the use of coherences (Ordinary Coherence (OC) between two signals, or Partial Coherence (PC) for multivariate signals) or Magnitude Squared Coherence (MSC) was introduced [Bendat 1986b, Brazier 1968]. Coherence allows to measure spatial correlations between signals in different frequency bands [Pfurtscheller 1999a]. This complex quantity¹ is sensitive to both changes in amplitude and in phase statistical relationships. As many other quantities, it must be generally computed on a sliding time interval (most often of constant duration) if the signals to be analyzed are not jointly stationary. In other words, if either the amplitude statistical dependence or the phase statistical dependence between the two signals under study changes, the coherence value is affected. If it exists a linear and time invariant relation between the two signals, the coherence magnitude is unity [Piersol 1993]. The coherence does not give any direct information on the existence of a true causal linear filter linking the two signals, but only information on an equivalent linear relationship.

2.1.2.2.2.2. Nonlinear techniques

Relatively to linear methods, nonlinear methods are designed to provide complementary information. Nonlinear neural time series analysis is motivated by the fact that many crucial neural processes have nonlinear characteristics (e.g. the regulation of voltage-gated ion channels corresponds to a steep nonlinear step-function relating membrane potential to current flow) and hence that most of the mechanisms at the origin of EEG/EMG signals are not strictly linear.

Influenced by a lot of applications in physics, a first family of methods based on nonlinear dynamical systems and deterministic chaos [Lorenz 1963] has been introduced to analyze neuronal signals. In the early 1980s, the concept of synchronization was introduced to measure neural connectivity. Synchronization is based on interacting chaotic oscillators [Pecora 1990, Pikovsky 1984]. It may be understood as an adjustment of rhythms of oscillating objects due to their weak interaction [Pikovsky 2003b]. In neuroscience studies, synchronization is mainly represented by the two following concepts – Phase Synchronization (PS) and Generalized Synchronization (GS) – [Ansari-Asl 2006, Boccaletti 2006, Chávez 2005, Chávez 2006b, Chávez 2007, Pfurtscheller 1999b, Pikovsky 2003a, Sakkalis 2009, Senhadji 2009]. Phase synchronization methods [Bhattacharya 2001, Chávez 2006a, Lachaux 1999, Rosenblum 2004, Rudrauf 2006] firstly extract an instantaneous phase from each signal and then quantify the degree of dispersion in the difference between estimated instantaneous phase using an appropriate index. One representative method capable to obtain a statistical measure

¹ To be rigorous, "coherency" is the normalized cross-spectrum at a given frequency and "coherence" is the magnitude or modulus of the coherency. Now, in the following, by abuse of terminology, we also use "coherence" for the complex value.

of the strength of PS in different areas of the brain is the Phase Locking Value (PLV) [Pereda 2005, Sakkalis 2009]. Generalized synchronization methods [Arnhold 1999, Stam 2002b] also include two steps; in the first one, state space trajectories are reconstructed from scalar time series signals and, in the second one, an index of similarity is computed to quantify the similarity (*i.e.* the existence of an instantaneous deterministic mapping between the two trajectories). PS and GS are also presented as helpful to analyze and/or explain mechanisms in certain diseases, such as the genesis of epileptic phenomena and Alzheimer's disease [Ansari-Asl 2007, Knyazeva 2010, Kramer 2007, Niedermeyer 1999, Senhadji 2009, Stam 2002a].

A second family of methods is based on Shannon information theory and the most representative among them is Mutual Information (MI) which measures the mutual dependence between two random variables by quantifying the amount of information gained about one over the other. As well known, it is a dimensionless quantity with units of bits, and can be thought of as the mean reduction in uncertainty about one random variable given knowledge of another. High mutual information indicates a large reduction in uncertainty and low mutual information a small reduction. A zero mutual information between two random variables sets means that the two sets are statistically independent [Cover 1991, Shannon 1949]. Beside mutual information, that can be applied to two sets of random variables, Partial Mutual Information (PMI) represents the information between two sets of random variables that is not contained in a third set of contextual random variables [Frenzel 2007].

2.1.2.2.3. Data processing in effective connectivity

Effective connectivity is defined as the direct or indirect influence that one neural system exerts over another, as well as the unidirectional or bidirectional influence between two neural systems [Horwitz 2003]. An illustration is given in Fig. 2.3 to show all possible influences among three neural systems. It describes the dynamic directional interactions among brain regions.

The interpretation of effective connectivity is that neural assemblies dynamically interact and eventually synchronize in local or distant regions either to accomplish perceptual, motor or cognitive functions [Delorme 2002], or as a consequence of pathological states such as those caused by epilepsy. In cognitive processes, such functions reflect complex interactions that include anticipation of the stimulus, attention to the stimulus and preparation for its associated actions [Zervakis 2011]. Such an interaction process can be realized through bidirectional or unidirectional coupling. The former case resembles mutual synchronization, where both systems adjust their rhythms to each other, whereas the latter case reflects causal interaction between the driver (initiating external force) and the response (the driven system). Mutual synchronization between two systems A and B can also be interpreted as a double causality, when the state evolution of each system is explicitly (in some admissible modeling) a function of its own present state and of the state of the other system. Effective connectivity can be based on a model eliciting causal links from structural connectivity knowledge or can be studied directly from the signals through concepts such as Granger causality or directed information, *i.e.* data-driven [Sakkalis 2011]. These techniques are discussed hereafter.

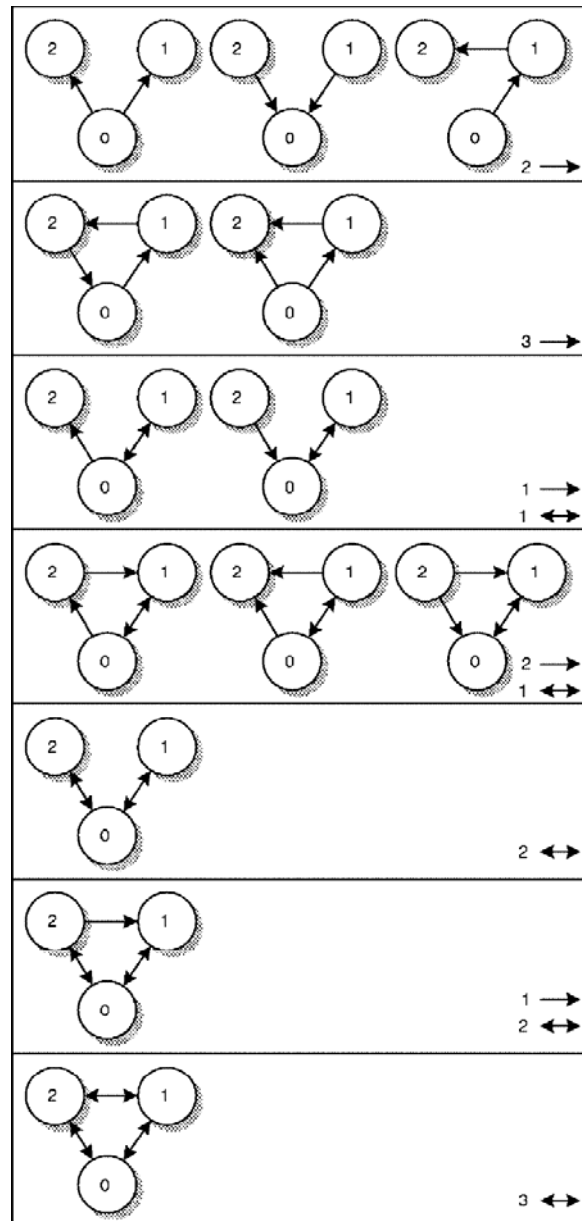


Fig. 2.3 – Possible influences between three populations (direct (e.g. from “0” to “2” in the first scenario of the first line) or indirect (e.g. from “0” to “2” in the third scenario of the first line), unidirectional (single-sided arrow) or bidirectional (doubled-sided arrow) influences) [Wendling 2000].

2.1.2.2.3.1. Model-based effective connectivity techniques

Neurobiological evidence and plausible theories generated from this evidence can form theoretical models that describe how brain areas interact and influence each other. This idea is the basis for model-based effective connectivity. Using this approach, innovative neurobiological models as well as hypotheses can be evaluated and causal interactions can be assessed.

Within the context of a specific brain state being investigated, neurophysiological data are considered as depending on independent brain sources activities that are spatially (and temporally) correlated. This idea dates back to the mid-1980s [Scherg 1985a,1985b], but more recently novel approaches have been developed to detect the local neural networks (brain sources) that underlie the spatial and temporal patterns that appear in raw EEG and MEG data [Friston 2010, Grech 2008,

Penny 2010]. In 1994, a technique called "covariance structural equation modeling" proposed by McIntosh and Gonzalez-Lima [McIntosh 1994] assigned effective connection strengths to anatomical pathways considered *a priori* (and integrated in a linear statistical model which is a static model, *i.e.* discarding the time delays as exogenous variables) that best match experimental covariances computed from data recorded during a (repeated) given task. A dynamical and nonlinear generalization of this approach called "Dynamic Causal Modeling (DCM)" [Friston 2003] operates in a Bayesian framework to estimate and make inferences about directed influences between variables. The key to DCM technique is that the response of a dynamic system can be modeled by a network of discrete but interacting neuronal sources described in terms of neural-mass [Jansen 1995, Lopes Da Silva 1974, Moran 2007] or conductance-based models [Morris 1981]. It was firstly introduced for fMRI [Friston 2009a, Stephan 2008], and later extended to EEG and MEG [Daunizeau 2009, Kiebel 2009].

At a higher level, and with a view to characterize the capacity of a system to integrate information, another approach to identify highly interactive brain regions and their directional interactions involves the use of a so-called effective information. Effective information between two subsets A and B, obtained by partitioning a given system S, is equal to mutual information between A and B when substituting node variables in A with statistically independent variables, hence maximizing entropy from A (under constrained variances). The maximization of this entropy allows capturing potential maximum causal influence from A to B [Tononi 2003] (the same procedure can be applied to test the influence from B to A). This method is very interesting at a conceptual level, but needs to manipulate an *a priori* known graph-structured Gaussian model of activities by removing selected links and computing again theoretical entropies. It allows essentially computing indices of causality between sub-systems on the basis of a pre-existing model as a structural linear model.

2.1.2.2.3.2. *Data-driven (model free) effective connectivity techniques*

In contrast to model-based techniques, "strictly" data-driven techniques do not assume any specific underlying physiological model or prior knowledge concerning spatial or temporal underlying relationships.

Wiener-Granger Causality (WGC) is one of the prototypical data-driven effective connectivity techniques [Bressler 2011]. In 1956 Norbert Wiener firstly recognized the importance of temporal ordering in the inference of causal relations and then introduced the notion that one variable (or time series) could be called "causal" to another if the ability to predict the second variable is improved by incorporating information about the first one [Wiener 1956]. Wiener however did not provide a practical (algorithmic) implementation of his idea. Such an implementation was introduced in 1969 by Clive Granger [Granger 1969] in the context of linear autoregressive (AR) models of stochastic processes used in econometrics. WGC reflects the extent to which a process X is leading another process Y , defined as an incremental predictability. By definition, a process X is said to Granger cause another process Y when the variance of the error in forecasting future values of Y , using an (optimal) forecast based on the observed values of both X and Y , is strictly smaller than the variance of the prediction error, using a (an optimal) forecast only based on the observed values of Y . So, if a variable can be predicted by the past information from a second variable and its own past information

better than from its own past information alone, then the second variable can be considered as "causal" to the first variable. Because the values of a variable at one time are predicted by values of other variables at earlier times, it is often said that WGC depends on "temporal precedence" (it does not include instantaneous causality). However, it is not sufficient that events in the other variables temporally precede similar events in the first variable. For WGC to be significant, statistically significant predictability must be established. In other words, non-zero values for WGC can usually be computed from any set of time series, but these values are meaningless unless it is determined that they are statistically significant [Bressler 2011]. WGC is intrinsically a time domain concept. But, with the introduction of AR causal models, the frequency decomposition of this fundamental tool was given later first by Saito and Harashima [Saito 1981] who proposed a spectral method called Directed COHerence (DCOH). Then, in [Geweke 1982] was introduced Frequency Geweke Causality (FGC). Geweke's work enabled the analysis of coupling between EEG frequency bands that have a well-known biomedical significance. With the development of WGC and FGC, a generalization of these concepts from bivariate to multivariate signals has been considered not only in the time domain [Gourévitch 2006, Hosoya 2001] but also in the frequency domain [Chen 2006, Ding 2006, Geweke 1984]. In recent years, these measures have also been extended to nonlinear cases [Barnett 2009, Chen 2004, Dhamala 2008, Fan 1995, Gourévitch 2006, Guo 2008, Hiemstra 1994, Marinazzo 2008] and largely applied in neuroscience [Bressler 2007, Chen 2006, Dhamala 2008, Ding 2006, Gourévitch 2006, Haufe 2011, Hesse 2003, Kocsis 2006, Quiroga 2000, Wang 2007, Winterhalder 2005]. In their method Saito and Harashima proposed to introduce an additional common source of noise to the autoregressive exogenous (ARX) model of Granger to take into account possible external sources influencing two time series. This common source of noise can be used to model the zero-delay coupling between these two series. In neurophysiology, this modified model was considered as more realistic than the instantaneous linear feedback measure proposed by Geweke [Geweke 1982]. In the subsequent decade, the DCOH has been widely applied in the neurophysiological system [Schnider 1989, Takigawa 1988, Takigawa 1996]. However, for the DCOH measure, only two channels are taken into account. The extension of the formalism to a greater number of channels poses difficulties rapidly increasing with their number. In order to solve this problem and determine the directional influences between the components in a multivariate system by a full multivariate frequency-domain based method, Kamiński and Blinowska [Kamiński 1991] introduced the Directed Transfer Function (DTF). The usefulness of the DTF method has been demonstrated in many studies (e.g. [Blinowska 2004, Eichler 2006, Kamiński 2001, Kamiński 2005, Veeramani 2004]), and it has been applied, for instance, to determine LFP propagation between brain structures of animals in different behavioural states [Korzeniewska 1997], to investigate EEG activity propagation in different sleep stages [Kamiński 1997], to localize epileptic loci [Fraszczuk 1998], and to study epileptogenesis [Medvedev 1999]. As a modification/generalization of the DCOH method and its variant, the DTF, the Partial Directed Coherence (PDC) was introduced by Baccalá and Sameshima [Baccalá 2001] to provide a clearer and more precise (or detailed) frequency domain connectivity image of Granger causality than that induced by directed coherence, especially for the simultaneous analysis of more than two time series ($N > 2$) [Sameshima 1999]. PDC is equivalent to DTF when

applied in the bivariate case but, in the multivariate case, PDC can detect direct pathways linking interacting brain regions and also discriminate between indirect and direct linking, which is not the case for DCOH and DTF [Astolfi 2006, Gourévitch 2006, Haufe 2011, Jung 2012, Popescu 2011, Schelter 2006a, Schelter 2006b, Vélez-Pérez 2008, Winterhalder 2005].

In the foregoing paragraph, only the magnitudes of all the mentioned measures are usually investigated. However, in certain diseases (e.g. epilepsy or Alzheimer's disease), some phase quantification can be found as significant in causality connectivity. With regard to this situation, Nolte *et al.* [Nolte 2008] proposed a measure, namely the Phase Slope Index (PSI), to robustly detect the direction of information flow and applied it in complex physical systems and to EEG/MEG signals [Haufe 2010, Haufe 2011, Nolte 2008, Nolte 2009, Nolte 2010]. The PSI basic hypothesis relies on the exploitation of the phase monotony between signals which appears when the frequency components of one signal precede temporally those of another signal. This method, based on the fact that pure delay between two signals implies linear phase in cross power spectral density, estimates the direction of propagation by computing the slope of the phase of the ordinary coherence function. The theoretical idea of this index is to represent properly relative time delays between spectral components of the two signals only in the frequency bands where the coherence is significant. Now, when we have to decide if two time series display direct or indirect relations, PSI which is based on the ordinary coherence function fails to distinguish between these two types of relations.

Although the above-mentioned measures have been extensively applied to deal with effective connectivity in neuroscience and obtained a series of encouraging significant triumph, when focusing on more complicated neurophysiological system, approaches in the scope of information theory are also considered to search for effective connectivity. One measure based on information theory is the so-called Transfer Entropy (TE) proposed by Schreiber in 2000 [Schreiber 2000] that has been widely used to estimate effective connectivity in neuroscience in the past decade [Besserve 2011, Chávez 2003, Kleeman 2011, Lindner 2011, Martini 2011, Neymotin 2011, Sabesan 2007, Vicente 2011, Wibral 2011a, Wibral 2011b]. As discussed in section 2.1.2.2.2, MI has been widely used to quantify the amount of common information between two systems. Unfortunately, MI neither contains dynamical nor directional information (if it is not computed for different delays). According to the fact that it shares some of the desired properties of mutual information but also takes the dynamics of information transport into account, TE is able to differentiate effectively driving and responding elements as well as to detect asymmetry in the subsystems' interaction. Since the direct application of this method as shown in [Schreiber 2000] may give unexpected results, Sabesan *et al.* focused their attention on internal parameters involved in TE to significantly improve its accuracy and robustness [Sabesan 2003, Sabesan 2007, Sabesan 2009a, Sabesan 2009b]. The selection of two crucial parameters has been discussed in [Sabesan 2009a, Sabesan 2009b]. The first one is the radius r that controls the size of the kernel used to estimate the multi-dimensional joint transitional probabilities. The second one pertains to the orders of the coupled Markov processes. Beside these approaches, another nonlinear technique using a nonlinear correlation coefficient [Lopes Da Silva 1989] has been applied to identify the dependency between two signals in the field of EEG analysis [Louis Dorr 2007, Wendling 2010].

2.1.3. Relations between structural, functional and effective connectivity

This section is devoted to summarize present research on high level information in neural networks and to the relationship between anatomical, functional and effective connectivity in the cortex, which represents nowadays a significant challenge to theoretical neuroscience. The main question to answer is to know how functional and effective brain networks emerge from structural brain connectivity.

This question naturally introduces the well-established notions of segregation and integration [Friston 1994, Tononi 1994]. At first glance, these notions seem opposite. Segregation refers to the existence of specialized neurons and brain areas, organized into distinct neuronal populations and grouped together to form segregated cortical areas. Functional segregation requires mutual independency in the firing of these specialized groups of neurons. The complementary principle, integration, gives rise to the coordinated activation of distributed neuronal populations and functional integration is based on the high coherence between their joint activities. The interplay of segregation and integration in brain networks generates information that is simultaneously highly diversified and highly integrated, thus creating patterns of high complexity [Sporns 2000].

A major fact is that application of network analysis techniques allows the comparison of brain connectivity patterns obtained from structural and functional studies. Most of modifications in cortical networks in the adult brain are observed on long timescales (associated with aging, disease progression, ...), whereas patterns of functional connectivity between brain areas undergo rapid fluctuations (induced by sensory input or cognitive tasks) on very short timescales (hundreds of milliseconds). These perturbations do not affect the stability of the global topological structure [Bassett 2006b, Valencia 2008]. Given this, the discovery of small-world attributes in functional connectivity patterns derived from fMRI, EEG and MEG studies raises the question how closely functional connections map onto structural connections. On the one hand, structural connection patterns are indeed major constraints for the dynamics of cortical circuits and systems, which are captured by functional and effective connectivity, and, on the other hand, in addition to the constraining influence of structural connections, rapid temporal fluctuations in functional or effective connectivity may reflect additional changes in physiological variables without necessarily modifying the structural brain connectivity. Computational models offer a complementary method to investigate structure-function relations in brain networks. Simulation studies of large-scale cortical networks demonstrated the emergence of complex spatio-temporal structures in neural correlations at multiple timescales [Honey 2007]. When functional connectivity is estimated from long time samples, the topology of structural and functional networks is identical. When functional networks are estimated on shorter time samples, the functional connectivity is less strongly constrained by the structural wiring diagram. If it is clear that structural networks have some impact on functional networks over long time periods, it is not so obvious how the structural topology participates in the reconfiguration of functional networks and how it is remodeled by the functional plasticity on a slower timescale. Future work must involve the parallel analysis of structural connectivity maps of the human brain and of patterns of functional and effective connectivity recorded in various conditions of rest or cognitive activation.

2.2. Neurocomputational models for epilepsy

As indicated before, an epileptic seizure corresponds to a "transient of signs and/or symptoms due to abnormal excessive or synchronous neuronal activity in the brain regions" [Fisher 2005]. During seizures, a variety of motor, sensory, cognitive or behavioral signs and symptoms are observed. Understanding the emergence of this disorder and predicting seizures are therefore great endeavors that have attracted a lot of attention from different scientific domains. This research leads to study the collective behavior of neurons in epileptic networks, and today bundles of experimental observations on epileptogenic networks still need a systematic integration to a large set of computational and analytical models [Touboul 2011].

Although the exact mechanisms leading to the various forms of epilepsy are still not well known, it is commonly admitted that the nature of the interactions between neurons and the properties of neurons themselves are altered in epileptogenic networks. Most studies suggest that epilepsy can often be linked with hyperexcitability and hypersynchronization (see [Babb 1989, Muñoz 2007, Noebels 1996]) of the involved networks. The investigation of such alterations is essential for the understanding of this condition, and could have important implications in future treatments.

For example, a recent and particularly powerful approach towards the understanding of seizure dynamics is the combination of electrophysiology with high-resolution fluorescent imaging [Sheth 2009]. But seizure mechanisms are too complex to understand without incorporating such measurements and observations into computational models. Only theoretical predictions can be quantitative in nature and allow direct comparison with experiments as access to various variables and parameters becomes increasingly available. Computational models including structural constraints, non observable state variables (bio-physiological model) and observable variables (direct problem model) are the most suitable tools to tie the advances made at various levels in epilepsy. Indeed, epileptic seizure is an example of a phenomenon which presumably cannot be properly understood without introducing computational representations to test physiological hypothesis [Mitra 2008].

There is a vast and valuable literature on physiologically relevant computational models which aim to link neuronal activity in more or less large neurons assemblies to EEG (depth or surface) observations. As in other computational biology areas, we are faced with the necessity to do compromise between complexity and capacity of capturing relevant mechanisms at the right scale(s). Models have been proposed from ionic channels scale to global brain electrical activity at the largest scale [Nunez 1995].

We mainly present here two classes of epilepsy computational models: (i) microscopic (biophysical network) models that describe the assembly of neurons at the cell level, and (ii) macroscopic (mean field) models that aim at capturing with a minimum number of equations the effective behavior emerging from the network activity. In the experimental part of this thesis only macroscopic models are used but a short presentation of microscopic models appears useful to give some prospects. Macroscopic models are an illustration of a modeling methodology in common use to aggregate variables introduced at a given representation scale to obtain more global structural and state variables linked with a small number of equations. This allows for reducing computational time

and makes easier the theoretical analysis. An example of this approach corresponds to the neuronal assemblies models of the cerebral cortex, referred also as "neural mass models" [Touboul 2011] in epilepsy.

2.2.1. Microscopic models

Sufficiently detailed biophysical network models of epilepsy are necessary to investigate the role of single neuron (role of biophysical and molecular properties) and neuronal networks topologies in causing seizure-like patterns of activity. For example, how can the shift in synaptic link or change in conductance of given ionic currents influence seizure-like brain activities? Firing patterns of individual neurons are mostly controlled by various ion channel conductance, synaptic inputs and the nearby microenvironment of neurons. Thus, these models look into the molecular bases of epilepsy and can suggest therapeutics based on their predictions. Despite 1) the limitations caused by the constraints of computational power and 2) the uncertainty of details of neuronal connections and intra neuronal biophysical properties, these models are useful areas of computational epilepsy. The general framework for these models is to reproduce the experimental data (essentially from electric or magnetic fields sensors but also from other modalities such as imaging) when adjusting intra neuronal and neural networks parameters and then to investigate the effect of various factors on the behavior of the networks. Such detailed neuronal models may provide an "access" to factors that are clearly usually inaccessible through only experimental means. These models start from a focus on a single synapse to networks composed of millions of neurons [E.P.F.L.]. In fact, the computing "bottleneck" of simulations' complexity can lead to the modeling of only few cells, or even of only one cell, if detailed intra-cellular processes are of major interest.

Network models with few cells may be useful for making predictions that could be generalized to larger networks. For example, Skinner *et al.* [Skinner 2005a, Skinner 2005b] presented a two-cell network model and examined the behavior of this network for certain sets of parameters. Then, they used the same set of parameters in a larger network of several tens of neurons and observed that the large network followed the same pattern of activity as seen in the smaller network. They extracted various parameters, such as synaptic and input currents relevant for the epileptic behavior under consideration.

In most models of epileptic seizures, a shift from dominant (or balanced) inhibition to dominant excitation in a neuronal network is considered to be responsible for the network transition from the preictal state to ictal state. In such imbalanced conditions, moderate perturbations can drive a neuronal network from physiological to seizure-like activity. A number of studies have addressed this point as can be seen in [Trevelyan 2006]. In [Van Drongelen 2005, Van Drongelen 2007], a neocortical network consisting of 656 neurons (512 principal cells (PCs) and 144 interneurons (INs)) exhibited seizure-like behavior when the synaptic excitation was decreased which constitutes a somewhat unforeseen result. To produce a transition to a seizure-like state the excitation to both excitatory cells and inhibitory cells needed to be reduced, while reducing each one alone was insufficient to cause the network to seizure-like behavior. These results illustrate the interest of

computational modeling in epilepsy, as intuition (in absence of any modeling) would have suggested that increased excitation alone could cause seizures.

Generally, works on microscopic network modeling have been done to represent a local neural activity in a localized region in cortex, assumed to present some isotropic organization. Besides, microscopic models for interdependent activity between two or more regions, hence introducing a form of structural anisotropy, have also been considered. For example, Destexhe *et al.* [Destexhe 1996] have done extensive work to explain absence seizures by developing detailed models of thalamocortical networks. A major finding is the role played by the thalamus and cortical regions in generating absence seizures. The typical spike-and-wave patterns of absence seizures require an interplay between both thalamus and cortex. Although seizures could be generated intracortically, the thalamus seems to be necessary for absence seizures. Spike-and-wave seizures disappear following the inactivation of the thalamus. Blocking GABA_B also leads to suppression of spike-and-wave seizures. These findings were experimentally confirmed by Bal *et al.* [Bal 2000].

In various seizure models, very fast oscillations having frequencies greater than 70 Hz have been observed, immediately before spontaneous seizures both *in vivo* and *in vitro* (see for example [Traub 2001] and [Worrell 2004]). Traub *et al.* performed detailed network modeling in conjunction with experimental studies to understand the mechanism of generation of very fast oscillations and also their transition to seizures [Traub 1982, Traub 2001, Traub 2003, Traub 2005] which can be of importance in the location of the epileptic focus.

2.2.2. Macroscopic models

On the one hand, the brain is a very complex nonlinear system and, on the other hand, it is difficult to build mathematical models requiring billions of state variables and parameters. Aiming at reproducing the global activity of the network that can be recorded by extracellular electrodes (LFP or EEG recording), neuroscientists have developed macroscopic models that can be derived through the use of mean-field limits [Ermentrout 1998, Faugeras 2009], and are built upon the facts that neurons are organized in different homogeneous populations sharing common characteristics (at least in a statistical sense).

These macroscopic models can be traced back to the 1970s with the typically extensions of the pioneering Wilson-Cowan equations [Wilson 1972,1973] who laid the theoretical foundations of these models, and drew upon the results of Mountcastle [Mountcastle 1957], and Hubel and Wiesel [Hubel 1963,1965,1968] who provided the first physiological evidence for the existence of macroscopic populations. Then, it has been progressively used to model the cat's olfactory system [Freeman 1973, 1975, 1987], or the alpha rhythm in the dog's EEG [Lopes Da Silva 1974, Lopes Da Silva 1976], and has now firmly demonstrated its ability to capture the dynamics of cortical areas [Suffczynski 2001, Suffczynski 2006, Wendling 2000, Wendling 2002, Zetterberg 1978].

In these models, the LFP or EEG signals reflect the global activity emerging from the microscopic interactions between thousands of neurons. These coupled activities are summarized through the interaction of macroscopic variables characterizing the mean activity of interconnected

neuronal sub-populations (mainly the pyramidal cells (PCs) and the interneurons (INs)). This point of view differs from the detailed biophysical models in the sense that it emphasizes the properties of neuronal populations as a whole instead of those of individual cells [Touboul 2011]. A particular important macroscopic variable in these models is the population's firing rate (averaged over all neurons of a population) that summarizes the spiking activity of neurons in a given subpopulation in detailed biophysical models. This variable is generally assumed to satisfy a nonlinear stochastic ordinary differential equation. The stochastic input is introduced to represent the stimulation of the considered population by other populations.

The simplest macroscopic models are the two subpopulations models (excitatory and inhibitory) with both inhibitory and excitatory connections among them (see Fig. 2.4) [Jansen 1993, Jansen 1995], while more complex models with more than two subpopulations are needed to produce complex EEG patterns [Lopes Da Silva 1974, Wendling 2000]. To model the activity of populations, two kinds of transfer functions are introduced: the first one corresponds to the dynamical linear pulse to wave transfer function which transforms the action potential average firing rate of the presynaptic population to the average inhibition or excitation membrane potential at the postsynaptic population, and the second one is the nonlinear static (represented by a sigmoid function) wave to pulse transfer function which transforms the wave activity (average level of membrane potential) of a subpopulation into the average firing rate of action potentials of the same subpopulation. Each subpopulation gives rise to two coupled differential equations of order one (details can be seen in [Wendling 2000]). This model was used by Wendling and colleagues to investigate Mesial Temporal Lobe Epilepsy (MTLE) [Wendling 2000, Wendling 2002, Wendling 2005]. In this kind of situation, multiple coupled populations can be constructed by coupling the single population model to other populations. This model exhibits various EEG patterns according to the ratio of excitatory to inhibitory inputs. Specifically, the transition from preictal to ictal state occurs when the ratio of excitation and inhibition increases above a certain threshold, thus supporting the hypothesis of increased relative excitation for seizure generation.

Further, models with more than two subpopulations, as that proposed by Wendling *et al.* in 2002 [Wendling 2002], are employed to explore more complex EEG patterns. This is specifically the kind of model we used in this thesis (see chapter 3 for details). In this model, Wendling *et al.* added a third subpopulation considering that various types of inhibitory projections to PCs could be impaired differentially in MTL epileptic hippocampus (Fig. 2.4 B). One INs subpopulation (dendritic projecting INs) projects to principal cell dendrites while the second INs subpopulation (basket cells) projects to PCs soma. The model indeed shows that transition from normal to fast ictal activity occurs when the two inhibition types (represented respectively by parameter B for slow inhibition and G for fast inhibition in Fig. 2.4) are reduced differentially (Fig. 2.4 C1). For moderately elevated excitation (represented by the parameter A in Fig. 2.4), the network switches from normal background activity to rhythmic spikes when slow dendritic inhibition is reduced (b-1 to b-2) while keeping the fast somatic inhibition fixed. Low-voltage rapid discharges are observed when slow inhibition is reduced further (b-2 to b-3). Finally, high amplitude paroxysmal activity is observed when slow dendritic inhibition is slightly increased and the fast somatic inhibition is reduced (b-3 to b-4). Representative trajectories from

regions (b1-b4) are shown in Fig. 2.4 C2. Results from the model compares well with clinical EEG data (Fig. 2.4 C3).

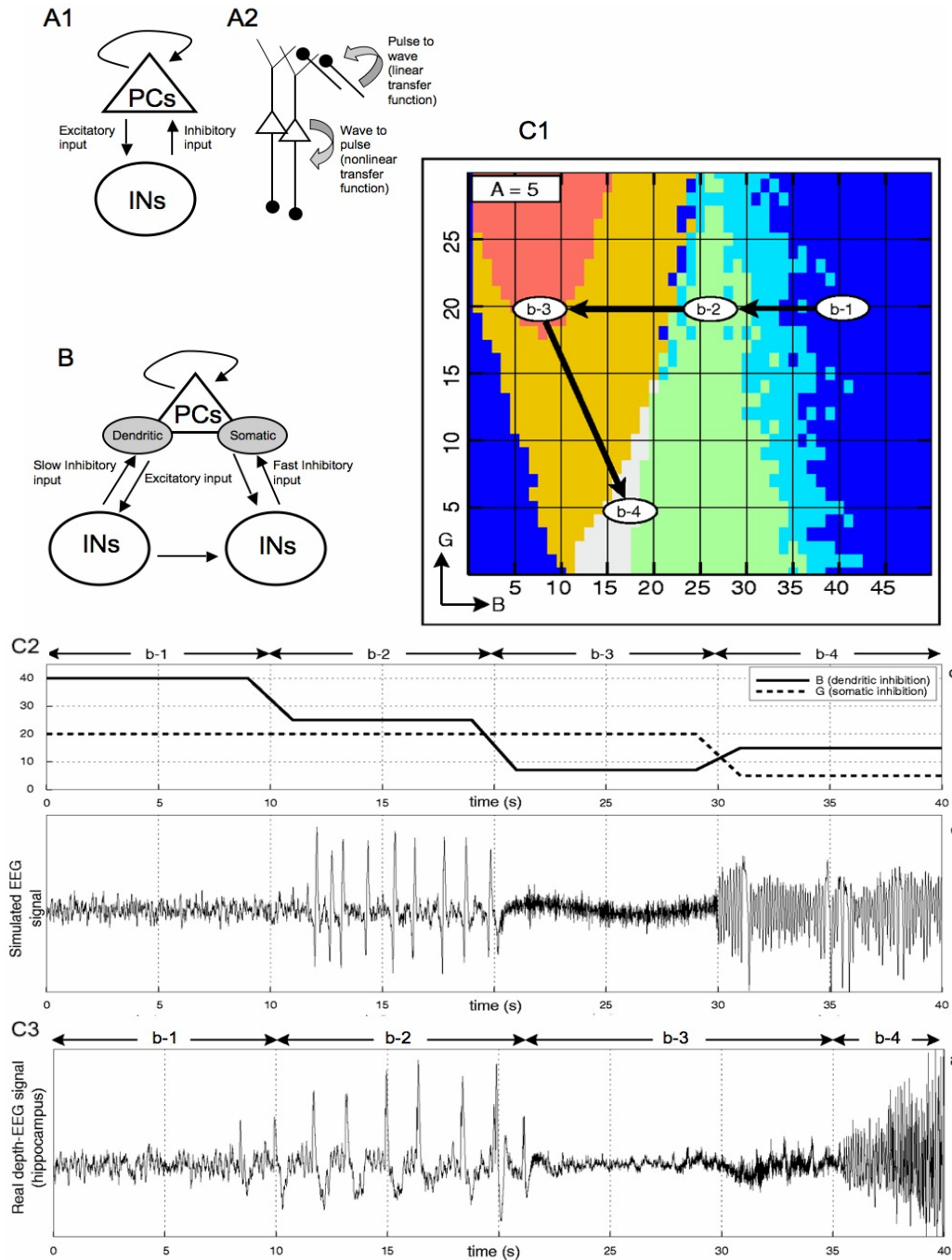


Fig. 2.4 - Basic structure of the mean field neuronal models. (A1) Two subpopulation model: PCs receive inhibitory inputs from INs and excitatory input from other PCs in the same subpopulation, while INs receive only excitatory input from PCs. Each subpopulation is characterized by two transfer functions - a linear pulse to wave and nonlinear wave to pulse (A2). The differential change in the MTLE hippocampus is taken into account by considering two interneuronal subpopulations – one projecting slow inhibition to the dendritic part of PCs and another projecting to the somatic part of PCs (B). Various behaviors exhibited by the model are summarized in (C1), where the slow dendritic inhibition is taken along the horizontal axis and fast somatic inhibition along the vertical axis. A representative trajectory from the model is shown in (C2), bottom panel, and the two inhibitory inputs in the top panel. The model compares well with the experimental data (C3) [Wendling 2002].

Considering four subpopulations, another approach was used by Suffczynski *et al.* [Suffczynski 2004] to develop a model for absence seizures that built on the model from Lopes da Silva *et al.* [Lopes Da Silva 1974]. These subpopulations are (1) cortical pyramidal cells (PY), (2) cortical inhibitory INs, (3) thalamocortical cells (TC) and (4) reticular thalamic cells (RE) (Fig. 2.5). The model consists of two modules, a cortical one and a thalamic one, that are mutually interconnected. The main results of this study show that (i) paroxysmal discharges represent bifurcations that occur in a neuronal network with bistability properties. This means that two stable states co-exist, one corresponding to the "normal on-going" EEG activity and the other to the paroxysmal oscillations, and that the system may undergo transitions from one state to another, (ii) the distributions of lengths of paroxysmal and "normal on-going" epochs are exponential, indicating that transitions between these two stable states occur randomly over time and that the probabilities for the transition between both states can be defined, (iii) probabilities of transitions between "normal on-going" neuronal activity and paroxysmal oscillations depend on a number of model parameters, (iv) paroxysmal oscillations can be annihilated by a well-timed pulse, (v) since random fluctuations in control parameters and/or dynamic variables can lead to the sudden onset of large amplitude paroxysmal activity, the occurrence of this type of phenomenon is unpredictable. A review of these population models is given in [Lopes Da Silva 2003].

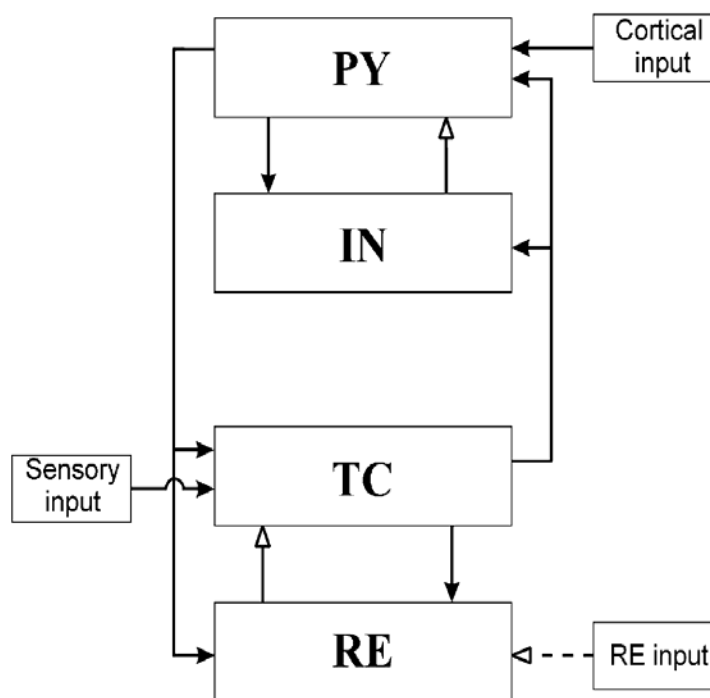


Fig. 2.5 - Schematic structure of connections in the thalamo-cortical network model consisting of cortical and thalamic modules.

The cortical module consists of two interconnected populations of PY and IN neurons.

The thalamic module consists of two interconnected populations of the TC and RE neurons.

TC cells project to both the PY and IN cells, while PY cells project to both the TC and RE cells.

The PY population receives external cortical excitatory input.

The TC population receives external sensory input.

The RE population receives external inhibitory input [Suffczynski 2004].

Recently, Kramer *et al.* [Kramer 2005, Kramer 2006] used the mean field approach to develop control strategies for human cortical electrical activity. They modeled the cortex as a system of fourteen differential equations, two for mean excitatory and inhibitory membrane potentials of cortical populations, and twelve equations mimicking the dynamics of excitatory and inhibitory synaptic and external inputs from other layers. For certain parameter sets, they found bifurcations between steady state (normal activity) and limit cycle (seizure-like activity) as a function of the excitation in the system. To eliminate the limit cycle and prevent seizures, the authors explored three control strategies and incorporated feedback controllers in the model dynamics – a linear controller, a differential controller and a filter controller – for controlling the large amplitude, stable "seizure" oscillations. These controllers appeared successful in terminating the oscillations in the model. One of them, the filter controller, was the most successful since other controllers pushed the model to a depolarized state while eliminating the limit cycle.

The mean field models have certain advantages over the more detailed models described in section 2.2.1. The usefulness of lumped models makes no doubt in the sense that macroelectrodes used for EEG recordings represent the average LFP arising from neuronal populations and so these models are suitable for exploring EEG activity from epileptic patients and looking into transitions from interictal to ictal states. Naturally, these models must be coherent with the aggregation of microscopic ones to return the true activity of the brain and, when building correctly these lumped models, the dynamical features of neuronal activity are not lost. Very often, epileptic activity spreads over quite extended regions and involves several cortical and sub-cortical structures. In such a context, these models are easy to analyze numerically because relatively few variables and parameters are involved. Hence, since the mean field models remain relatively simple, they represent the best alternative to describe physiologically epileptic processes occurring in "large-scale" systems. On the other hand, the main drawback of such macroscopic models is that they fail to suggest molecular and cellular mechanisms of epileptogenesis whereas the detailed biophysical network models are best suited for understanding the molecular and cellular bases of epilepsy and thus are well positioned to suggest therapeutics that could target molecular pathways. Now, although biophysically explicit modeling is the primary technique to look into the role played by experimentally inaccessible variables in epilepsy, the usefulness of detailed biophysical models is limited by constraints in computational power, uncertainties in detailed knowledge of neuronal systems, and the required simplification for the numerical analysis. Unlike the lumped models, detailed network models are much more difficult to analyze numerically for a range of parameters as their dynamics take place in many dimensional state space. An intermediate "across-scale" approach, establishing relationships between sub-cellular/cellular variables of detailed models and "aggregated" parameters governing macroscopic models, would be a very useful strategy to cover the gaps between these two modeling approaches.

Chapter 3

Experimental protocol

We present in the next chapter different manners to compute a connectivity index. Given the number of indices we proposed added to those of the literature, the compared evaluation of performance appears already a difficult task. In a standard way, we used simulated signals and we also tested our algorithms on a small amount of real signals recorded on a guinea-pig (animal epilepsy model). As largely introduced in chapter 2, the practical use of these connectivity indices is to answer the following questions for a pair of signals X_i and X_j ^{II}:

- Is there a statistical link between X_i and X_j , the negative answer corresponding to the H_0 hypothesis?
- If yes, can we decide that X_i influences X_j (H_{ij} hypothesis) or that X_j influences X_i (H_{ji} hypothesis) or that influence is reciprocal (H_{i-j} hypothesis)?

The first point corresponds to the existence or non-existence of a functional connectivity. The second point corresponds to characterizing this functional connectivity in terms of effective connectivity, given that we ignore if this connectivity is due to a direct structural (anatomical) link between the sub-systems generating the two signals.

The questions remain the same conditionally on a third signal X_k (or conditionally on an ensemble of auxiliary signals):

- Conditionally to X_k , is there a statistical link between X_i and X_j , the negative answer corresponding to the $H_{0/k}$ hypothesis?
- If yes, and conditionally to X_k , does X_i influence X_j ($H_{ij/k}$ hypothesis), or does X_j influence X_i ($H_{ji/k}$ hypothesis), or do we have a reciprocal influence ($H_{i-j/k}$ hypothesis)?

^{II} From now on, a capital letter, e.g. X , represents a stochastic process (or variable) and a small letter, e.g. x , is its realization.

The first point relates to the question of the existence of a direct functional connectivity and the second one to the question of the existence of a direct effective connectivity. Note that the term "direct" refers relatively to the context defined by X_k (or to an ensemble of auxiliary signals).

H_0 and $H_{0/k}$ hypotheses are generic (*i.e.* they only refer respectively to the notions of ordinary and conditional independence). The others are more subtle and require the definition of the notion of influence, which can be considered in the sense of Wiener-Granger (see chapter 4) and which require in practice to identify a generic model (or at least to indicate the orders of some underlying Markov processes). Note that, if the effective influence, in a direction or in the opposite one, is recognized in the Wiener sense, the question of the existence of instantaneous connectivity is not addressed.

Simulated models proposed in the following have been chosen so that the responses are non ambiguous (ground-truth). This is effectively the case for AR models (linear or nonlinear) and also for models of neuronal populations (called hereafter physiological models), since their intrinsic equations explicitly code the effective connectivity links. Note that the physiological models used in this work are of the same nature (in mathematical and semantic points of view) as the dynamical models used in DCM approach, but the present use of such models is to get more realistic simulated signals than those generated by AR modeling. A standard DCM procedure would estimate coupling coefficients comparable to those described in section 3.1.3. Let us indicate two problems in this approach. The first one comes from the fact that it answers to connectivity only by values of these coefficients and neither on properties of the resulting dynamics nor on the strength of the causality coupling. If a null coupling parameter implies an absence of causality, we have to evaluate the strength of the dynamic link when the value of the parameter differs from zero. From this point of view, signal-dependent based approaches directly address the right question. The second problem relies on the other parameters such as intra-population parameters. As a matter of fact, in the field of epilepsy, parameters linked to excitation and inhibition are dependent on the type of activity and must be also estimated which is not so trivial [Frogerais 2008].

In summary, to judge the relevance of our tools, it was necessary to get a ground-truth and, beyond standard AR models, to test plausible physiological models. All these models which represent instantiations of various *a priori* propagation graphs are presented in sections 3.1.1, 3.1.2 and 3.1.3.

Finally, the response to the veracity of some H_T hypothesis, relative to the interpretation of observations $O = (X_i, X_j)$ or $O = (X_i, X_j, X_k)$, implies the introduction of a statistical hypothesis test of the form: H_T is accepted if a statistics $T(O)$ of real values belongs or not to a defined domain of acceptance (generally an interval of \mathbb{R}). For the hypotheses relative to the ordinary (resp. conditional) connectivity considered in this work, $T(O)$ takes respectively the form $Ind_{X_i \rightarrow X_j}$ (resp. $Ind_{X_i \rightarrow X_j / X_k}$) where Ind represents one of the indices proposed in chapter 4. Details concerning the strategy and the afferent statistics are given in section 3.2.2.

3.1. Database

3.1.1. Linear autoregressive models

The generic linear autoregressive model we considered is as follows:

$$\begin{cases} x_1(t) = 0.95\sqrt{2}x_1(t-1) - 0.9025x_1(t-2) + w_1(t) \\ x_2(t) = -0.5x_1(t-1) + 0.25\sqrt{2}x_2(t-1) - \beta x_3(t-3) + w_2(t) \\ x_3(t) = -\alpha x_1(t-2) - 0.5x_2(t-2) - 0.25\sqrt{2}x_3(t-2) + w_3(t) \end{cases} \quad (3.1)$$

where $w_j(t)$, $j = 1, 2, 3$, are realizations of independent white noises W_j with zero mean and unit variance. The parameter β ($\beta = 0.5$) is introduced to model bidirectional flow between signals X_2 and X_3 . The introduction of the parameter α allows to consider two patterns of causal interactions, either direct relations ($\alpha = 0.5$) or indirect relations ($\alpha = 0$). These 4 combinations of parameters lead to 4 models of connectivity (model 1: $\alpha = 0$, $\beta = 0$, model 2: $\alpha = 0.5$, $\beta = 0$, model 3: $\alpha = 0$, $\beta = 0.5$, model 4: $\alpha = 0.5$, $\beta = 0.5$) (see Fig. 3.1). The power spectral densities (PSD) of these signals can be found in Appendix A.

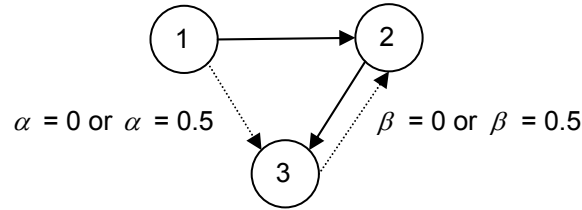


Fig. 3.1 - Simulated model to generate three signals X_1 , X_2 and X_3 . Integer i in a circle corresponds to X_i .

3.1.2. Nonlinear autoregressive models

The stochastic model for the nonlinear signals is governed by the following equations:

$$\begin{cases} x_1(t) = 3.4x_1(t-1)(1-x_1^2(t-1))e^{-x_1^2(t-1)} + w_1(t) \\ x_2(t) = 3.4x_2(t-1)(1-x_2^2(t-1))e^{-x_2^2(t-1)} - 0.5x_1(t-1) + 0.25\sqrt{2}x_2(t-1) - \beta x_3(t-3) + w_2(t) \\ x_3(t) = 3.4x_3(t-1)(1-x_3^2(t-1))e^{-x_3^2(t-1)} - \alpha x_1(t-2) - 0.5x_2(t-2) - 0.25\sqrt{2}x_3(t-2) + w_3(t) \end{cases} \quad (3.2)$$

where $w_j(t)$, $j = 1, 2, 3$, are realizations of independent white noises W_j with zero mean and unit variance. As previously, the parameter β is introduced to model bidirectional flow between signals X_2 and X_3 (see Fig. 3.1). In this case, this parameter is set to 0.5. The parameter α is introduced to take into account two patterns of causal interactions, either direct relations ($\alpha = 0.5$) or indirect relations ($\alpha = 0$). As previously, we get 4 models (model 1: $\alpha = 0$, $\beta = 0$, model 2: $\alpha = 0.5$, $\beta = 0$, model 3: $\alpha = 0$, $\beta = 0.5$, model 4: $\alpha = 0.5$, $\beta = 0.5$) (see Fig. 3.1). The power spectral densities of these signals can be found in Appendix A.

3.1.3. Physiological models

For the physiology based model, we use a time continuous SDE (Stochastic Differential Equations) model simulated in discrete time to represent the electrical activity of distant, and possibly coupled, neuronal populations. It is based on the physiology and was introduced in [Wendling 2005] to produce outputs which can be interpreted as intracranial electroencephalographic signals similar to those recorded with proximal electrodes in hippocampus.

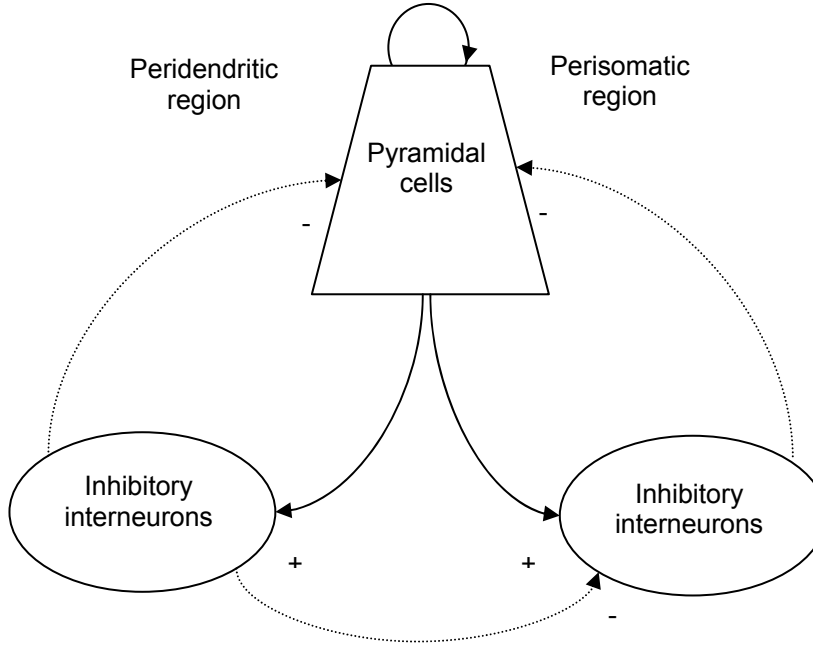


Fig. 3.2 - Interactions between different neuronal subpopulations of the hippocampus [Wendling 2005].

For each population, the model generates a local mean field activity (mean membrane potential) that is converted to an intracerebral electroencephalographic (iEEG) signal recorded in a proximal electrode using a quasi-static transfer function [Wendling 2005]. Each population consists of three subpopulations that mutually interact: a major subpopulation of excitatory neurons and two subpopulations of inhibitory neurons (Fig. 3.2). The main excitatory subpopulation has a feedback action on itself. The cells of the first inhibitory subpopulation correspond to the interneurons that project to the dendritic region of the main population. Regarding the second subpopulation, it is composed of inhibitory interneurons for which the synaptic contacts are carried out in the somatic area of primary neurons.

The corresponding mathematical description of this model is given in graphical form in Fig. 3.3 [Frogerais 2008]. In Fig. 3.3, the three subpopulations P_e , P_{s_i} and P_{f_i} appear in three boxes underlined by dotted lines, corresponding respectively to the main subpopulation and to the two other subpopulations with slow and fast inhibitory interneurons. The input $W(t)$ summarizes the influence of distant afferent neurons and is represented by the formal derivative of a Brownian process (*i.e.* white Gaussian noise). The impulse response $G_{PH}h_{PH}$ (where G_{PH} is a multiplicative constant) is that of an instrumentation high-pass filter, whose output is sampled at 256 Hz. Its transfer function is

- $h_e(t)$: impulse response for the excitation,
 $h_{fi}(t)$: impulse response for the somatic inhibition (fast inhibition),
 $h_{si}(t)$: impulse response for the dendritic inhibition (slow inhibition).

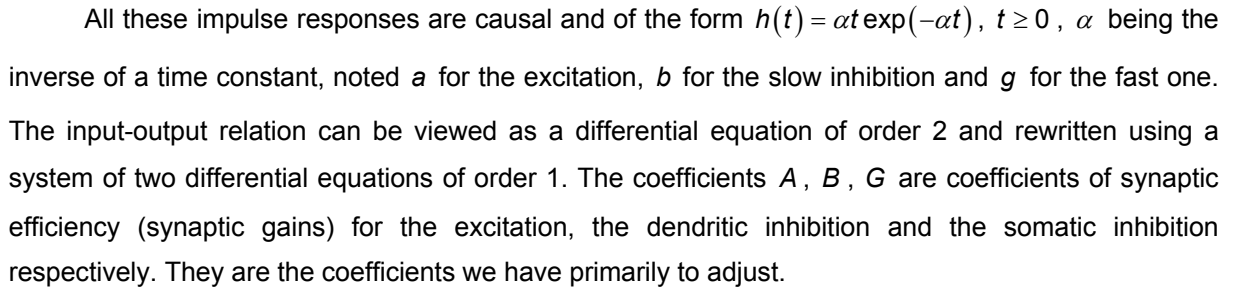


Fig. 3.3 - Block diagram of the population model.

The corresponding equations of the system are:

This system is of order 11. As a matter of fact, we have 7 transfer functions of order 2 for the pre-somatic filters (see Fig. 3.3), which introduce 14 state variables. However, the three transfer functions Ah_e placed respectively after the coefficients C_1 , C_3 and C_5 can be merged into a single transfer function by switching them with the three coefficients to save four state variables [Wendling 2008]. The transfer function of the output high-pass filter is of order 1 and requires only one additional state component which therefore brings to 11 ($14 - 4 + 1$) the dimension of the state vector and that of the differential system. Only parameters A , B and G are supposed to vary during a transition from a normal process to the epileptic seizure. The connectivity constants C_i , $i = 1, \dots, 7$, and the synaptic time constants $1/a$, $1/b$ and $1/g$ and the parameters of the input Gaussian process are supposed to be known and fixed as in [Wendling 2001, Wendling 2005]. These values are reported in Tab. 3.1. For this set of data, the generated signals and their PSD are shown in Fig. 3.4.

Synaptic time constants		
$1/a$	Excitatory	1/100
$1/b$	Slow inhibitory	1/30
$1/g$	Fast inhibitory	1/350
Connectivity constants		
C_1	$P_e - P_e$	135
C_2	$P_e - P_e$	108
C_3	$P_e - P_{s_i}$	33.8
C_4	$P_{s_i} - P_e$	33.8
C_5	$P_e - P_{f_i}$	40.5
C_6	$P_{s_i} - P_{f_i}$	13.5
C_7	$P_{f_i} - P_e$	121.5
White Gaussian noise (input)		
m_p	mean	90
σ	diffusion	30
Sigmoid		
e_0		$2.5s^{-1}$
v_0		6mV
r		$0.56mV^{-1}$

Tab. 3.1 - Example of model constants of hippocampus.

Since the pyramidal cells are excitatory neurons that project their axons to other areas of the brain, the model accounts for this organization by using the average pulse rate of action potentials from the main cells of one population i as an excitatory input to the main cells inputs of a second

population j . In addition, this connection from population i to j is represented by a parameter K^{ij} which is proportional to the number of corresponding active axonal links for a given type of cerebral activity. An appropriate setting of this parameter allows for building systems where the neuronal populations are coupled either unidirectionally or bidirectionally. The other parameters of this model are internal parameters (inside the population itself). They include excitatory and inhibitory gains in the feedback loops as well as coefficients related to the number of synaptic contacts between subpopulations. These parameters are adjusted to control the intrinsic activity of each population (normal background activity versus epileptic activity).

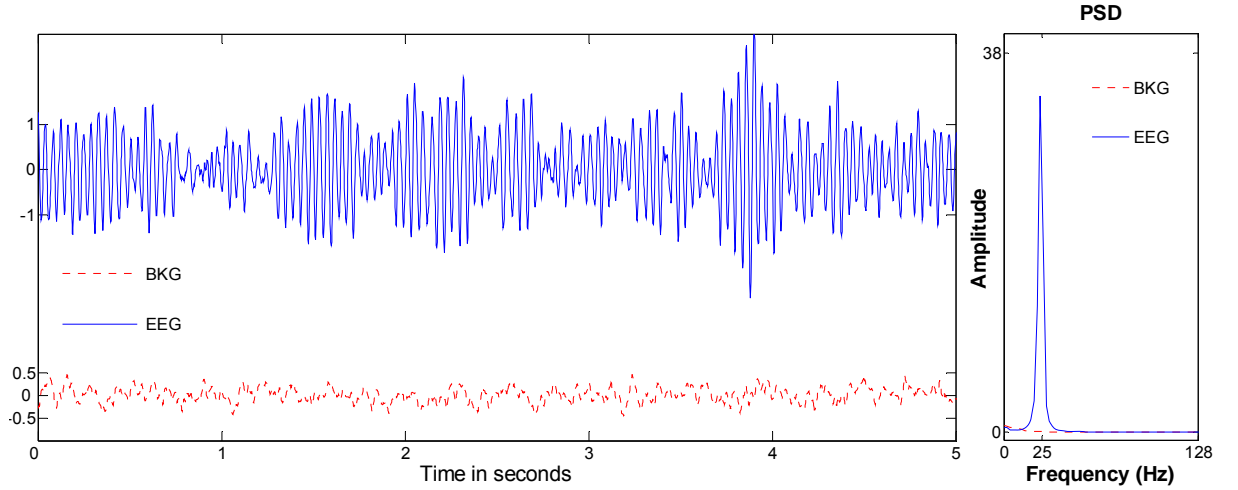


Fig. 3.4 - Example of signals generated by the model.

Left panel: simulated background signal (red dashed line, $A = 5$, $B = 1$ and $G = 20$) and simulated EEG signal (blue solid line, $A = 5$, $B = 3$ and $G = 20$).

Right panel: corresponding PSD (BKG: background activity, EEG: epileptic activity).

For the four scenarios we define hereafter, the values of the parameters A , B , G and g can be found in Appendix B, and the PSD of the signals generated in these scenarios in Appendix A.

3.1.3.1. First scenario: Model 1

For this scenario, we have only one parameter K , with $K = K^{12} = K^{23}$, varying from 0 to 1500 by step of 500. For $K \neq 0$, the three populations are epileptic and, for $K = 0$, populations 2 and 3 become normal background activity.

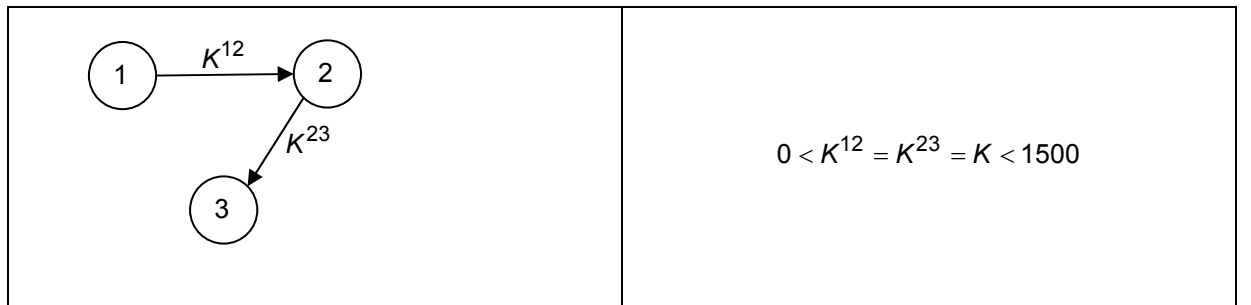


Fig. 3.5 - Population 1 drives population 2 and population 2 drives population 3. The parameter K varies from 0 to 1500 by step of 500.

3.1.3.2. Second scenario: Model 2

For the second scenario, population 1 drives population 2, and both populations 1 and 2 drive the 3rd one; different values of the parameters K^{ij} are chosen. A first choice consists in setting $K^{12} = 1500$ and $K^{13} = K^{23} = 1500$, and the second choice consists in setting $K^{12} = 1500$ and $K^{13} = K^{23} = 750$.

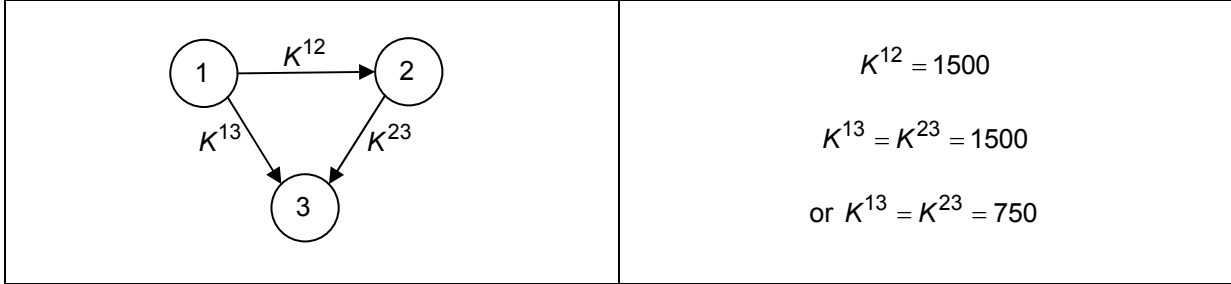


Fig. 3.6 - Population 1 drives population 2, and these two populations drive population 3.

3.1.3.3. Third scenario: Model 3

For the third scenario, population 1 still drives population 2, and populations 2 and 3 are bi-directional coupled populations. We choose the following parameters $K^{12} = K^{32} = K^{23} = 1500$.

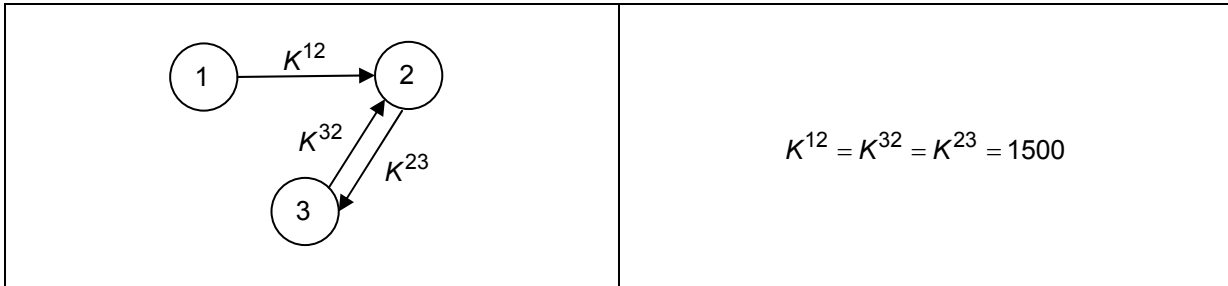


Fig. 3.7 - Population 1 drives population 2 and populations 2 and 3 interact.
The three populations are epileptic.

3.1.3.4. Fourth scenario: Model 4

For the fourth scenario, population 1 drives populations 2 and 3, and these two populations interact jointly. The values retained for the parameters are: $K^{12} = K^{13} = K^{32} = K^{23} = 1500$.

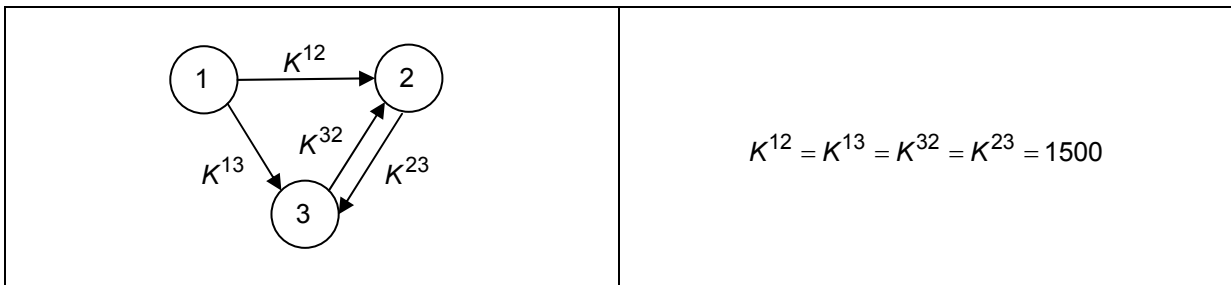


Fig. 3.8 - Population 1 drives populations 2 and 3.
Populations 2 and 3 interact through the introduction of parameters K^{23} and K^{32} .
Parameters are such that the three populations are epileptic.

3.1.4. Real signals

The interictal and ictal epileptiform activities are induced by a 3-min arterial perfusion of 50 μm bicuculline in the isolated guinea pig brain preparation maintained *in vitro* (more details can be seen in [Uva 2005]). Simultaneous extracellular recordings are performed in the piriform cortex, in the medial and lateral entorhinal cortex, in area CA1 of the hippocampus, and in the perirhinal cortex. The position of the five recording electrodes is illustrated in Fig. 3.9 and the corresponding recorded EEG signals are also plotted on this figure. In Chapter 5, some phases of these signals, and more specifically the FOA (Fast Onset Activity) corresponding to the time interval [30 s; 40 s], are tested by the measures we developed.

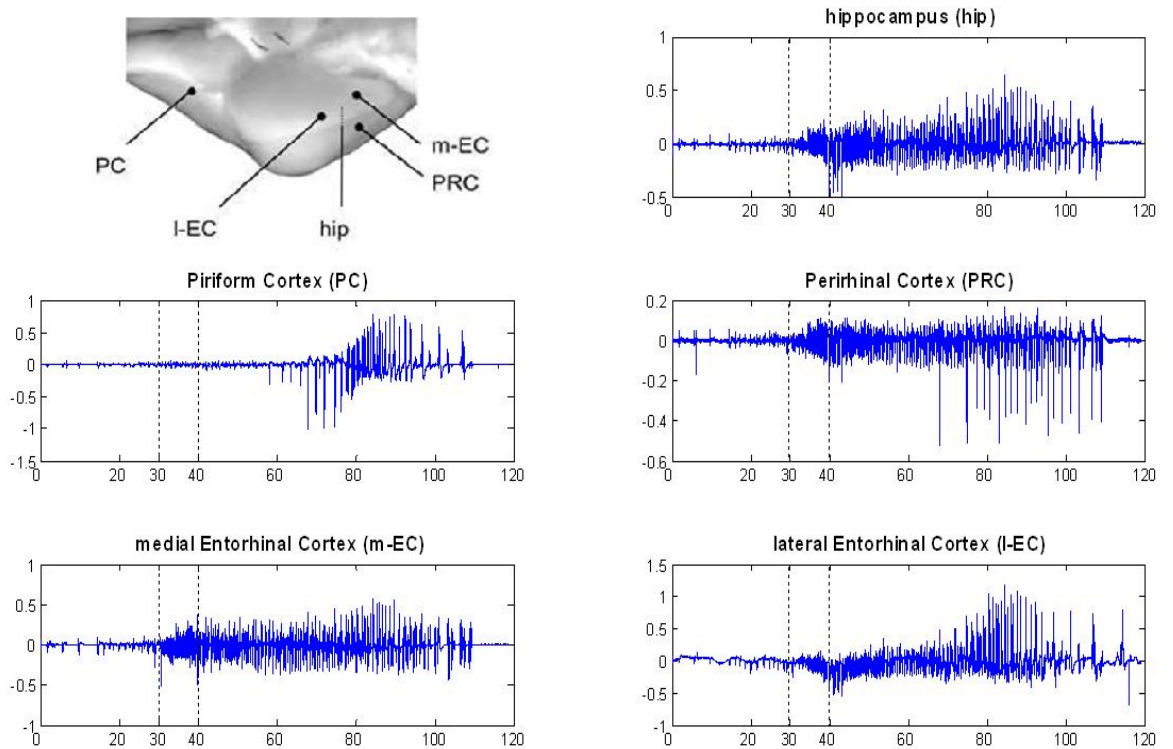


Fig. 3.9 - Typical pattern of interictal-to-ictal transition induced by a 3-min arterial perfusion of 50 μm bicuculline in the isolated guinea pig preparation maintained *in vitro*. Simultaneous extracellular recordings performed in the piriform cortex, in the medial and lateral entorhinal cortex, in area CA1 of the hippocampus, and in the perirhinal cortex. The position of the recording electrodes is illustrated in the left upper panel. The corresponding recorded EEG signals are also plotted.

3.2. Methodology of evaluation

3.2.1. Determination of graphs

In a first step, we compute indices of effective connectivity (see chapter 4) on the signals described in the previous section and then we propose to establish graphs of information flow thanks to the mean values of these indices. Beyond the visual information, the difficulty lies in the determination of a threshold to decide whether there is effective connectivity or not, and so we must think about the introduction of a test for each pair (X_i, X_j) (non conditional analysis) or each triplet (X_i, X_j, X_k) (conditional analysis).

All indices $Ind_{X_i \rightarrow X_j}$ and $Ind_{X_i \rightarrow X_j / X_k}$ being theoretically equal to zero under H_0 hypothesis and all the more important as the (functional or effective) link becomes stronger, the test will be of the following form in the two cases:

$$T_2(X_i, X_j) = Ind_{X_i \rightarrow X_j} > \lambda \Rightarrow H_0 \text{ rejected and } H_{ij} \text{ accepted}$$

$$T_3(X_i, X_j, X_k) = Ind_{X_i \rightarrow X_j / X_k} > \lambda \Rightarrow H_0 \text{ rejected and } H_{ij/k} \text{ accepted}$$

Given the expressions of T_2 and T_3 , we have to determine a threshold λ . Commonly, we have to choose λ such as $P(T_2 > \lambda / H_0) \leq pfa$ or $P(T_3 > \lambda / H_{0/k}) \leq pfa$, where pfa is an error probability of first kind (also called type I error or probability of false alarm or false negative probability) fixed by the user. Now, the theoretical computation of these quantities is only accessible in very simple cases where the law of T_2 or T_3 under the null hypothesis becomes invariant relatively to the statistics of the signals under study. In practice, it is generally necessary to use an empirical distribution of T_2 or T_3 using surrogate data as suggested next. For a given triplet of signals (i, j, k) , a propagation graph is obtained in the following manner:

For $(i, j) = (1, 2), (2, 1), (1, 3), (3, 1), (1, 2), (2, 1)$:

- 1) determine $\lambda_{i,j}$ (resp. $\lambda_{i,j/k}$)
- 2) compute $Ind_{X_i \rightarrow X_j}$ (resp. $Ind_{X_i \rightarrow X_j / X_k}$)
- 3) if $Ind_{X_i \rightarrow X_j} > \lambda_{ij}$, accept H_{ij} (resp. if $Ind_{X_i \rightarrow X_j / X_k} > \lambda_{ij/k}$, accept $H_{ij/k}$)
(otherwise reject the hypothesis)

end

The first step corresponds to the computation of an adaptive threshold according to the observed signals. Let us indicate that we can also obtain this threshold considering a preliminary step using a learning database.

3.2.2. Analysis of the distributions under H_0

For a set of three populations corresponding to signals x_1, x_2 and x_3 (realizations of random signals $X_i, i = 1, 2, 3$) for which we measure an index $Ind_{X_i \rightarrow X_j}$ from population i to population j (given that a model of connectivity is available for these three populations), the problem is to evaluate the deviation from the H_0 hypothesis (corresponding to independent signals X_i and X_j), since it is difficult to obtain a theoretical distribution of $Ind_{X_i \rightarrow X_j}$ under H_0 (resp. of $Ind_{X_i \rightarrow X_j / X_k}$ under $H_{0/k}$). This difficulty can be removed using surrogate data synthesized from the original data and guaranteeing their independence to get a reference statistics under H_0 (resp. $H_{0/k}$). To this end, we must develop a strategy to modify the realizations x_i or x_j so that the two signals are independent while preserving their frequential characteristics (which largely influence the variance of any statistics computed from these observations).

For this, we investigate two methods described hereafter.

1st method

Given two realizations x_i and x_j (coming from initially dependent signals X_i and X_j), we compute the Fourier transform of one of them, for example x_j , get the corresponding analytical signal, change its phase for each discrete frequency by a random value on the interval $[0, 2\pi]$ and, after inverse Fourier transform, take the real part of the signal. In this way, we get a signal X_j' independent of X_i , or at least decorrelated, and which has the same power spectral density as the original signal X_j [Theiler 1992]. This technique of random phase allows for generating decorrelated signals whose spectral marginal characteristics (at the second order) are preserved. To get a statistical distribution of $Ind_{X_i \rightarrow X_j}$ under H_0 , (resp. $Ind_{X_i \rightarrow X_j | X_k}$ under $H_{0/k}$), we repeat the procedure for a sufficient number of independent realizations x_j' (of X_j'). A threshold can be obtained by computing the quantile corresponding to the desired *pfa* probability. This strategy can be used either for simulated or real signals and allows for computing a threshold for a given pair of signals (computation of the threshold in the loop of the previous algorithm).

2nd method

In the first method, the destruction of the phase somewhat modifies the "nonlinear" temporal characteristics of the signals, which can be harmful. An alternative to build signals under H_0 (resp. $H_{0/k}$) hypothesis to avoid this drawback is described hereafter.

Let us give a series of M independent realizations (x_i^m, x_j^m, x_k^m) , $m = 1, \dots, M$, obtained with the same model (same structure, same parameters) or using a repeated controlled experience, for instance on an animal model. Suppose we want to test the relation between signals X_i and X_j . To build independent pairs (X_i, X_j') preserving marginal laws, we must choose a sufficient number of pairs $(X_i^m, X_j^{m'})$, $m \neq m'$, for statistical accuracy. In this second method, the determination of a threshold can be done outside the loop of the previous algorithm.

Therefore, to validate our approaches on simulated signals (linear and nonlinear AR models and physiological models), we decide to compute our connectivity indices under H_0 (resp. $H_{0/k}$) hypothesis according to one of the aforementioned methods. We restrict this study to a limited number of indices, i.e. those for which the "subjective" graphs appear coherent with the structures of the models used for simulations, and to only one type of connectivity (model 4) for the three kinds of simulated signals (linear and nonlinear AR models and physiology-based model). In chapter 5, for each explored technique, the results obtained under H_0 (resp. $H_{0/k}$) are compared with those obtained on the observations. Concerning the real signals, since we only tested a reduced database, we retained the approach based on surrogate data using random phase.

Chapter 4

Methods

As discussed earlier, our work aims to analyze intracerebral EEG (iEEG) signals recorded during epileptic seizures and particularly to identify the brain structures involved in the different phases of such seizures, quantify the information carried by the multiple observations and investigate as finely as possible the relationships between structures taking part in these epileptic events.

Investigating brain connectivity usually underlies the successive apprehension of the different stages of the model, *i.e.* its specification, identification and the causal inference [Valdes-Sosa 2011]. At this level, it should be noted that we are not concerned by the definition of specific models used to measure causal relationships between signals, even if our purpose is much more focused on effective connectivity than on functional connectivity. In other words, we focus on the determination of effective connectivity, which pertains to the generic detection of causal relationships between neural systems. Let us insist on the fact that causality is an epistemological concept that is particularly difficult to understand via equations. In this respect, the notion of causality can be thought as follows:

- It can be seen as based on a temporal precedence, *i.e.* causes precede their effects. Temporal precedence is taken into account in the notion of Granger causality which is a statistical concept based on prediction, generally referenced as WGC (Wiener-Granger Causality) [Bressler 2011], and also called WAGS (Wiener-Akaike-Granger-Schweder) influence by Valdes-Sosa in [Valdes-Sosa 2011];
- It can be seen as a physical influence, *i.e.* changing causes changes consequences. It underlies the notion of intervention and control and was first formalized in 2000 by Pearl [Pearl 2000]. Observing or estimating some activity at a network node provides information on potential effects to remote nodes. Operating physically on this activity removes any other influence the node receives.

This distinction is important because it is the basis of any statistical detection of a causal influence. In the context of brain connectivity, identifying causal relationships between two brain areas

can be thought either in terms of improving the prediction of temporally distinct neural events or in terms of estimating the remote effect of changes on the events.

In this chapter, first of all, we recall the well established concept of Wiener-Granger causality and we present different variants of the Wiener-Granger causality index in the time domain (WGCI-P and WGCI-C) and in the frequency domain (FGCI-P and FGCI-C). Then, we propose new indices based on a more recent measure, the phase slope index, where different coherence functions can be considered to detect and differentiate various patterns of causality relations (*i.e.* direct vs indirect and unidirectional vs bidirectional). Next, we study an information theoretic method, named Transfer Entropy (TE), and a new strategy identifying the model order is proposed to enhance the effectiveness and the robustness of this technique. Finally, since TE only deals with causal information by pairwise analysis, another approach, called Conditional Transfer Entropy (CTE), is considered to distinguish between direct and indirect causality interactions in trivariate time series. To help the reader, and due to the large number of proposed methods, we synthesized them according to their typology in Appendix C.

4.1. Granger causality

The basic idea gets back to Wiener [Wiener 1956] who considered that, if the prediction of a time series could be improved by incorporating the knowledge of a second one, then the second series is said to have a causal influence on the first one. Wiener's idea lacks the machinery for practical implementation. Later, Granger formalized the prediction idea in the context of linear regression models [Granger 1969]. According to Granger causality, if a signal X_1 "Granger-causes" a signal X_2 , then past values of X_1 should contain information that helps predict X_2 above and beyond the information contained in past values of X_2 alone. Granger causality was developed in 1960s and has been widely used in economics since the 1960s. However, it is only within the last few years that applications in neuroscience have become popular. Specifically, if the variance of the autoregressive prediction error of the first time series at the present time is reduced by including past measurements from the second time series, then the second time series is said to have a causal influence on the first one. The roles of the two time series can be reversed to address the question of causal influence in the opposite direction. The interaction discovered in this way may be reciprocal or it may be unidirectional. Two additional developments of Granger's causality idea are important. First of all, when three or more time series exist simultaneously, the causal relation between two of these series may be direct, mediated by a third one, or a combination of direct and indirect relations. This situation can be addressed by the technique of conditional Granger causality. Secondly, natural time series, evolving in economics or neurobiology, may display oscillatory aspects in specific frequency bands. In this case, it can be desirable to have a spectral representation of causal influence. A decomposition of the Granger causality in the frequency domain was suggested in 1980s by Geweke [Geweke 1982, 1984]. His measure of Granger causality was derived from the spectral representation of a bivariate autoregressive model. Several researchers [Chen 2006, Chicharro 2012, Ding 2006, Wang 2007] have subsequently worked on the spectral Granger causality measure of Geweke and proposed various testing procedures. These recent developments lead to a renewed interest in frequency-based

Granger causality approaches, in particular in neuroscience and fMRI-related fields (see [Baccalá 2001, Brovelli 2004, Chen 2006, Gourévitch 2006, Salazar 2004]) but also in economics [Nishiyama 1997, Winer 1986]. Most on-going research is based on the framework of Geweke. In this section, we reconsider the original framework and give the essential mathematical elements of Granger causality in time and frequency domains before proposing new developments.

4.1.1. Time domain

4.1.1.1. Bivariate case

Let X_1 and X_2 be two zero-mean signals whose time observations are noted $x_1(t)$ and $x_2(t)$, with $t = 1, 2, \dots, T$, where T is the signal length. If we model each observation $x_1(t)$ and $x_2(t)$ by a univariate AR model of order p , we have

$$x_1(t) = \sum_{k=1}^p \alpha_1(k) x_1(t-k) + u_1(t) \quad (4.1)$$

$$x_2(t) = \sum_{k=1}^p \alpha_2(k) x_2(t-k) + u_2(t) \quad (4.2)$$

where each signal, at time t , depends only on its own past, u_1 and u_2 are white Gaussian noise realizations. Now, if we model both signals $x_1(t)$ and $x_2(t)$ by a bivariate AR model of order p , we write

$$x_1(t) = \sum_{k=1}^p \alpha_{11}(k) x_1(t-k) + \sum_{k=1}^p \alpha_{12}(k) x_2(t-k) + w_1(t) \quad (4.3)$$

$$x_2(t) = \sum_{k=1}^p \alpha_{22}(k) x_2(t-k) + \sum_{k=1}^p \alpha_{21}(k) x_1(t-k) + w_2(t) \quad (4.4)$$

where each signal depends not only on its own past but also on the past of the second signal, and w_1 and w_2 are white Gaussian noise realizations. Let us begin with the case of two signals by studying the causality $X_1 \rightarrow X_2$. From a univariate model, the quality of the representation of X_2 may be evaluated from the variance of the prediction error $\Gamma_{X_2|X_2^-}$, where X_2^- symbolizes X_2 past. Using a bivariate model, the variance of the prediction error becomes $\Gamma_{X_2|X_2^-, X_1^-}$. If X_1 causes X_2 in the Wiener-Granger sense, then $\Gamma_{X_2|X_2^-, X_1^-}$ is smaller than $\Gamma_{X_2|X_2^-}$. Considering Pairwise analysis, the level of Wiener-Granger Causality Index (WGCI-P) from X_1 to X_2 is then evaluated by

$$\text{WGCI}_{X_1 \rightarrow X_2}^{\text{-P}} = \ln \frac{\Gamma_{X_2|X_2^-}}{\Gamma_{X_2|X_2^-, X_1^-}} = \ln \frac{\text{var}(U_2(t))}{\text{var}(W_2(t))}. \quad (4.5)$$

Reciprocally, the WGCI from X_2 to X_1 can be evaluated as:

$$\text{WGCI}_{X_2 \rightarrow X_1} - P = \ln \frac{\Gamma_{X_1|X_1^-}}{\Gamma_{X_1|X_1^-, X_2^-}} = \ln \frac{\text{var}(U_1(t))}{\text{var}(W_1(t))}. \quad (4.6)$$

4.1.1.2. Multivariate case

Modelling Q signals, X_1, X_2, \dots, X_Q , by a multivariate AR model of order p , we write

$$\begin{pmatrix} x_1(t) \\ \vdots \\ x_Q(t) \end{pmatrix} = \sum_{k=1}^p \begin{pmatrix} \alpha_{11}(k) & \alpha_{12}(k) & \dots & \dots & \alpha_{1Q}(k) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \alpha_{mn}(k) & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \alpha_{Q1}(k) & \dots & \dots & \dots & \alpha_{QQ}(k) \end{pmatrix} \begin{pmatrix} x_1(t-k) \\ \vdots \\ x_Q(t-k) \end{pmatrix} + \begin{pmatrix} w_1(t) \\ \vdots \\ w_Q(t) \end{pmatrix} \quad (4.7)$$

where w_m , $m = 1, 2, \dots, Q$, are realizations of white Gaussian noises. The coefficient $\alpha_{mn}(k)$ evaluates the linear interaction of $x_n(t-k)$ on $x_m(t)$, whatever m, n . These coefficients α_{mn} , $m, n \in \{1, \dots, Q\}$, are estimated by least squares method [Kariya 2004]. The model order p is commonly determined by Akaike's Information Criterion (AIC) [Akaike 1973] or the Bayesian Information Criterion (BIC) [Schwarz 1978] that are detailed in Appendix D.

In the case of multiple signals, we can analyze independently each pair of signals as previously. However, pairwise analysis in a multivariate case cannot distinguish between direct and indirect coupling as shown in Fig. 4.1 (a) and (b). In the multivariate case, to disambiguate such cases, we must deal with direct causality from X_i to X_j conditionally to other $Q-2$ signals and consider the Conditional Wiener-Granger Causality Index, noted hereafter $\text{WGCI}_{X_i \rightarrow X_j | X_{Q-2}^-}$ -C, and defined by Eq. (4.8) where the numerator is the variance of the prediction error of X_j by taking all signals into account except X_i

$$\text{WGCI}_{X_i \rightarrow X_j | X_{Q-2}^-} - C = \ln \frac{\Gamma_{X_j | X_1^- \dots X_{i-1}^- X_{i+1}^- \dots X_Q^-}}{\Gamma_{X_j | X_1^- \dots X_Q^-}}. \quad (4.8)$$

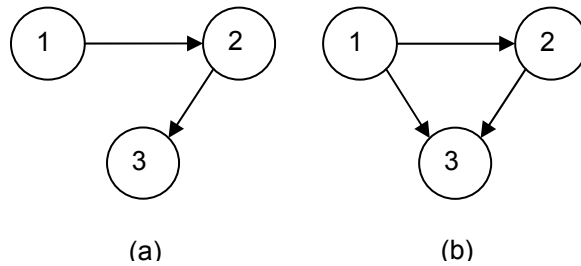


Fig. 4.1 – Two patterns of causal interactions.

(a) causality relation from signal X_1 to signal X_3 is indirect and mediated by signal X_2 .

(b) both direct and indirect causalities exist from signal X_1 to signal X_3 .

4.1.2. Frequency domain^{III}

By using Fourier methods it is also possible to examine Granger causality in the spectral domain. This can be very useful for neurophysiological signals, where frequency decompositions are often of interest. Intuitively, spectral Granger causality from X_1 to X_2 measures the fraction of the total power at frequency f of X_2 that is contributed by X_1 . In 1980s, Geweke [Geweke 1982, 1984] found a new technique for time series decomposition that expressed the Granger causality from the time domain in terms of its frequency content. Then, Ding *et al.* [Bressler 2007, Chen 2006, Ding 2006, Wang 2007] have further developed this concept and improved it. In the following, we present the formulations of this technique in the bivariate and trivariate cases respectively.

4.1.2.1. Bivariate case

Using the lag operator L , $L^k x_m(t) = x_m(t-k)$, with $m = 1, 2$ and $k = 1, 2, \dots, p$, we rewrite Eqs. (4.3) and (4.4) in terms of the lag operator

$$\begin{pmatrix} B_{11}(L) & B_{12}(L) \\ B_{21}(L) & B_{22}(L) \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} w_1(t) \\ w_2(t) \end{pmatrix} \quad (4.9)$$

where $B_{11}(L) = 1 - \sum_{k=1}^p \alpha_{11}(k)L^k$, $B_{22}(L) = 1 - \sum_{k=1}^p \alpha_{22}(k)L^k$, $B_{12}(L) = -\sum_{k=1}^p \alpha_{12}(k)L^k$, $B_{21}(L) = -\sum_{k=1}^p \alpha_{21}(k)L^k$, with $B_{11}(0) = 1$, $B_{12}(0) = 0$, $B_{21}(0) = 0$, $B_{22}(0) = 1$. Fourier transforming both sides of Eq. (4.9) leads to:

$$\underbrace{\begin{pmatrix} D_{11}(f) & D_{12}(f) \\ D_{21}(f) & D_{22}(f) \end{pmatrix}}_{D(f)} \begin{pmatrix} X_1(f) \\ X_2(f) \end{pmatrix} = \begin{pmatrix} W_1(f) \\ W_2(f) \end{pmatrix} \quad (4.10)$$

where the components of the coefficient matrix $D(f)$ are $D_{11}(f) = 1 - \sum_{k=1}^p \alpha_{11}(k)e^{-2i\pi fk}$, $D_{12}(f) = -\sum_{k=1}^p \alpha_{12}(k)e^{-2i\pi fk}$, $D_{21}(f) = -\sum_{k=1}^p \alpha_{21}(k)e^{-2i\pi fk}$ and $D_{22}(f) = 1 - \sum_{k=1}^p \alpha_{22}(k)e^{-2i\pi fk}$; $X_1(f)$, $X_2(f)$, $W_1(f)$ and $W_2(f)$ are respectively the "Fourier transforms" of $x_1(t)$, $x_2(t)$, $w_1(t)$ and $w_2(t)$.

Rewriting Eq. (4.10), we obtain

^{III} As far as we work in the frequency domain (in this section and following ones), a notation as $X(f)$ denotes the spectral representation of X without any distinction between the random process and a particular spectral realization. Moreover, and more rigorously, in case of a stationary random process X , it is well known that $X(f)$ does not exist as a classical Fourier transform and a representation with the infinitesimal random spectral components $dX(f)$ should be introduced. Now, for sake of clarity, we consider subsequently $X(f)$ as $dX(f)$.

$$\begin{pmatrix} X_1(f) \\ X_2(f) \end{pmatrix} = \underbrace{\begin{pmatrix} H_{11}(f) & H_{12}(f) \\ H_{21}(f) & H_{22}(f) \end{pmatrix}}_{H(f)} \begin{pmatrix} W_1(f) \\ W_2(f) \end{pmatrix} \quad (4.11)$$

where the transfer function is $H(f) = D^{-1}(f)$ whose components are $H_{11}(f) = \frac{1}{\det D} D_{22}(f)$, $H_{12}(f) = -\frac{1}{\det D} D_{12}(f)$, $H_{21}(f) = -\frac{1}{\det D} D_{21}(f)$, and $H_{22}(f) = \frac{1}{\det D} D_{11}(f)$.

Then, we can write the spectral matrix as follows

$$S(f) = \begin{pmatrix} S_{11}(f) & S_{12}(f) \\ S_{21}(f) & S_{22}(f) \end{pmatrix} = H(f) \Sigma H^\dagger(f) \quad (4.12)$$

where $S_{11}(f)$ and $S_{22}(f)$ are the auto-spectral density functions of signals X_1 and X_2 respectively, $S_{12}(f)$ and $S_{21}(f)$ are the cross-spectral density functions between signals X_1 and X_2 , $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} = \begin{pmatrix} \text{var}(W_1(t)) & \text{cov}(W_1(t), W_2(t)) \\ \text{cov}(W_2(t), W_1(t)) & \text{var}(W_2(t)) \end{pmatrix}$, $\text{var}(\cdot)$ stands for variance, $\text{cov}(\cdot)$ stands for covariance, and \dagger denotes Hermitian transpose.

To obtain a frequency decomposition of the causality, we first observe the auto spectrum of X_1 :

$$S_{11}(f) = H_{11}(f) \Sigma_{11} H_{11}^*(f) + 2 \Sigma_{12} \text{Re}(H_{11}(f) H_{12}^*(f)) + H_{12}(f) \Sigma_{22} H_{12}^*(f) \quad (4.13)$$

where $*$ denotes the conjugate value.

It is instructive to consider the case where $\Sigma_{12} = 0$. In this case, there is no instantaneous causality and the interdependence between X_1 and X_2 is entirely due to their interactions through the regression terms on the right hand sides of Eqs. (4.3) et (4.4). The spectrum of X_1 is composed of two terms. The first term, viewed as the intrinsic part, involves only the variance of W_1 , which is the noise term that drives the time series X_1 . The second term, viewed as the causal part, involves only the variance of W_2 , which is the noise term that drives X_2 . This power decomposition into an "intrinsic" term and a "causal" term is important to define a measure for spectral domain causality. When Σ_{12} is not zero, it becomes harder to distribute the power of X_1 to different sources. Here, we consider a transformation introduced by Geweke [Geweke 1982] that removes the cross term and makes the identification of an intrinsic power term and a causal power term possible. The procedure is called normalization and it consists of left-multiplying both sides of Eq. (4.10) by the following matrix

$$\tilde{P} = \begin{pmatrix} 1 & 0 \\ -\Sigma_{21} / \Sigma_{11} & 1 \end{pmatrix}. \quad (4.14)$$

It results in

$$\underbrace{\begin{pmatrix} D_{11}(f) & D_{12}(f) \\ \tilde{D}_{21}(f) & \tilde{D}_{22}(f) \end{pmatrix}}_{\tilde{D}(f)} \begin{pmatrix} X_1(f) \\ X_2(f) \end{pmatrix} = \begin{pmatrix} W_1(f) \\ \tilde{W}_2(f) \end{pmatrix} \quad (4.15)$$

where $\tilde{D}_{21}(f) = D_{21}(f) - \frac{\Sigma_{21}}{\Sigma_{11}} D_{11}(f)$, $\tilde{D}_{22}(f) = D_{22}(f) - \frac{\Sigma_{21}}{\Sigma_{11}} D_{12}(f)$, and $\tilde{W}_2(f) = W_2(f) - \frac{\Sigma_{21}}{\Sigma_{11}} W_1(f)$.

So, the new transfer function $\tilde{H}(f) = \tilde{D}^{-1}(f)$ is given by

$$\tilde{H}(f) = \begin{pmatrix} \tilde{H}_{11}(f) & \tilde{H}_{12}(f) \\ \tilde{H}_{21}(f) & \tilde{H}_{22}(f) \end{pmatrix} = \begin{pmatrix} H_{11}(f) + \frac{\Sigma_{21}}{\Sigma_{11}} H_{12}(f) & H_{12}(f) \\ H_{21}(f) + \frac{\Sigma_{21}}{\Sigma_{11}} H_{22}(f) & H_{22}(f) \end{pmatrix}. \quad (4.16)$$

By construction, it is easy to notice that the random spectral components $W_1(f)$ and $\tilde{W}_2(f)$ are uncorrelated, that is $\text{cov}(W_1(f), \tilde{W}_2(f)) = 0$. The variance of the normalized noise term of the second

equation in Eq. (4.9) is $\tilde{\Sigma}_{22} = \Sigma_{22} - \frac{\Sigma_{21}^2}{\Sigma_{11}}$. From Eq. (4.15), in the same way as we obtained Eq. (4.13),

the spectrum of X_1 is now:

$$S_{11}(f) = \tilde{H}_{11}(f) \Sigma_{11} \tilde{H}_{11}^*(f) + H_{12}(f) \tilde{\Sigma}_{22} H_{12}^*(f). \quad (4.17)$$

The first term is interpreted as the intrinsic power and the second one as the causal power of X_1 due to X_2 . This is an important relation because it explicitly identifies the portion of the total power of X_1 at frequency f that is contributed by X_2 . Based on this interpretation, we define Geweke's causal influence from X_2 to X_1 (by pairwise analysis) at frequency f as

$$\text{FGCI}_{X_2 \rightarrow X_1}(f) - P = \ln \frac{S_{11}(f)}{\tilde{H}_{11}(f) \Sigma_{11} \tilde{H}_{11}^*(f)}. \quad (4.18)$$

Note that this definition of causal influence is given in terms of the intrinsic power rather than in terms of the causal power. It is expressed in such a manner that the causal influence is zero when the causal power is zero (*i.e.* the intrinsic power equals the total power), and increases as the causal power increases (*i.e.* the intrinsic power decreases).

By taking the transformation matrix $\hat{P} = \begin{pmatrix} 1 & -\Sigma_{12}/\Sigma_{22} \\ 0 & 1 \end{pmatrix}$ and performing the same analysis as

previously (see Eqs. (4.14) and (4.15)), the new transfer function $\hat{H}(f)$ is

$$\hat{H}(f) = \begin{pmatrix} \hat{H}_{11}(f) & \hat{H}_{12}(f) \\ \hat{H}_{21}(f) & \hat{H}_{22}(f) \end{pmatrix} = \begin{pmatrix} H_{11}(f) & H_{12}(f) + \frac{\Sigma_{12}}{\Sigma_{22}} H_{11}(f) \\ H_{21}(f) & H_{22}(f) + \frac{\Sigma_{12}}{\Sigma_{22}} H_{21}(f) \end{pmatrix}. \quad (4.19)$$

The spectrum of $x_2(t)$ is found to be:

$$S_{22}(f) = \hat{H}_{21}(f) \hat{\Sigma}_{11} \hat{H}_{21}^*(f) + \hat{H}_{22}(f) \Sigma_{22} \hat{H}_{22}^*(f) \quad (4.20)$$

where $\hat{\Sigma}_{11} = \Sigma_{11} - \Sigma_{12}^2 / \Sigma_{22}$.

Then, we get the causal influence from X_1 to X_2 :

$$\text{FGCI}_{X_1 \rightarrow X_2}(f) - P = \ln \frac{S_{22}(f)}{\hat{H}_{22}(f) \Sigma_{22} \hat{H}_{22}^*(f)}. \quad (4.21)$$

4.1.2.2. Trivariate case

As already discussed in the time domain, Geweke's causality index in the frequency domain in pairwise analysis cannot disambiguate the two situations illustrated in Fig. 4.1. A second measure, namely the conditional Geweke causality index in the frequency domain, noted FGCI-C, is presented hereafter to solve this problem in the case of trivariate time series. In the following, the procedure is developed when considering the causality from signal X_2 to signal X_1 conditioned on the third signal X_3 , leading to the index $\text{FGCI}_{X_2 \rightarrow X_1 | X_3}(f) - C$; of course, the same procedure applies when we exchange the role of the three signals.

Consider three stochastic processes X_1 , X_2 , and X_3 . From Eq. (4.8) the conditional Wiener-Granger causality influence from X_2 to X_1 conditional on X_3 in the time domain is expressed by

$$\text{WGCI}_{X_2 \rightarrow X_1 | X_3} - C = \ln \frac{\Gamma_{X_1 | X_1^- X_3^-}}{\Gamma_{X_1 | X_1^- X_2^- X_3^-}}. \quad (4.22)$$

To derive the spectral decomposition of the time domain conditional Wiener-Granger causality in Eq. (4.22), a normalization procedure is performed (as in Eq. (4.15)) for the bivariate and trivariate cases.

Firstly, let us define the joint bivariate AR model of $x_1(t)$ and $x_3(t)$ as follows

$$\begin{pmatrix} \bar{B}_{11}(L) & \bar{B}_{13}(L) \\ \bar{B}_{31}(L) & \bar{B}_{33}(L) \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_3(t) \end{pmatrix} = \begin{pmatrix} u_1(t) \\ u_3(t) \end{pmatrix} \quad (4.23)$$

where $\bar{B}_{11}(0) = 1$, $\bar{B}_{33}(0) = 1$, $\bar{B}_{13}(0) = 0$, $\bar{B}_{31}(0) = 0$, and obviously $\Gamma_{X_1 | X_1^- X_3^-} = \text{var}(U_1(t))$.

Applying the normalization procedure achieved by using a transformation matrix

$$\begin{pmatrix} 1 & 0 \\ -\frac{\text{cov}(U_3(t), U_1(t))}{\text{var}(U_1(t))} & 1 \end{pmatrix} \text{ as in Eq. (4.14), the normalized equations are}$$

$$\begin{pmatrix} \tilde{B}_{11}(L) & \tilde{B}_{13}(L) \\ \tilde{B}_{31}(L) & \tilde{B}_{33}(L) \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_3(t) \end{pmatrix} = \begin{pmatrix} \psi_1(t) \\ \psi_3(t) \end{pmatrix} \quad (4.24)$$

so that the covariance between the corresponding processes $\Psi_1(t)$ and $\Psi_3(t)$ is imposed to be zero ($\text{cov}(\Psi_1(t), \Psi_3(t)) = 0$) to get the frequency decomposition considering a conditional dependence. In this case, $\tilde{B}_{11}(0) = 1$, $\tilde{B}_{33}(0) = 1$, $\tilde{B}_{13}(0) = 0$, and generally $\tilde{B}_{31}(0)$ is not zero [Geweke 1984].

The joint trivariate AR model used for deriving the frequency decomposition of the conditional dependence is

$$\begin{pmatrix} B_{11}(L) & B_{12}(L) & B_{13}(L) \\ B_{21}(L) & B_{22}(L) & B_{23}(L) \\ B_{31}(L) & B_{32}(L) & B_{33}(L) \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} = \begin{pmatrix} w_1(t) \\ w_2(t) \\ w_3(t) \end{pmatrix} \quad (4.25)$$

where the covariance matrix of the noise terms is

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} \end{pmatrix} = \begin{pmatrix} \text{var}(W_1(t)) & \text{cov}(W_1(t), W_2(t)) & \text{cov}(W_1(t), W_3(t)) \\ \text{cov}(W_2(t), W_1(t)) & \text{var}(W_2(t)) & \text{cov}(W_2(t), W_3(t)) \\ \text{cov}(W_3(t), W_1(t)) & \text{cov}(W_3(t), W_2(t)) & \text{var}(W_3(t)) \end{pmatrix}. \quad (4.26)$$

Obviously, $\Gamma_{x_1|x_1^-x_2^-x_3^-} = \text{var}(W_1(t))$. The normalization process involves left-multiplying both sides of Eq. (4.25) by the matrix P defined by

$$P = P_2 \cdot P_1 \quad (4.27)$$

$$\text{where } P_1 = \begin{pmatrix} 1 & 0 & 0 \\ -\Sigma_{21}\Sigma_{11}^{-1} & 1 & 0 \\ -\Sigma_{31}\Sigma_{11}^{-1} & 0 & 1 \end{pmatrix} \text{ and } P_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -(\Sigma_{32} - \Sigma_{31}\Sigma_{11}^{-1}\Sigma_{12}) (\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})^{-1} & 1 \end{pmatrix}.$$

Then, the normalized trivariate AR model can be written

$$\begin{pmatrix} \hat{B}_{11}(L) & \hat{B}_{12}(L) & \hat{B}_{13}(L) \\ \hat{B}_{21}(L) & \hat{B}_{22}(L) & \hat{B}_{23}(L) \\ \hat{B}_{31}(L) & \hat{B}_{32}(L) & \hat{B}_{33}(L) \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} = \begin{pmatrix} \varepsilon_1(t) \\ \varepsilon_2(t) \\ \varepsilon_3(t) \end{pmatrix} \quad (4.28)$$

where the covariance matrix of the noise terms is

$$\hat{\Sigma} = \begin{pmatrix} \hat{\Sigma}_{11} & 0 & 0 \\ 0 & \hat{\Sigma}_{22} & 0 \\ 0 & 0 & \hat{\Sigma}_{33} \end{pmatrix} = P \Sigma P^\dagger \quad (4.29)$$

where $\hat{\Sigma}_{33} = \Sigma_{33} - \Sigma_{31}\Sigma_{11}^{-1}\Sigma_{13} - (\Sigma_{32} - \Sigma_{31}\Sigma_{11}^{-1}\Sigma_{12})(\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})^{-1}(\Sigma_{23} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{13})$, $\hat{\Sigma}_{11} = \Sigma_{11}$, and $\hat{\Sigma}_{22} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$.

Based on the relations of different variances, Geweke derived the following important relation of the conditional causality in the time domain [Geweke 1984]

$$\text{WGCI}_{X_2 \rightarrow X_1 | X_3} = \text{WGCI}_{X_2 \Psi_3 \rightarrow \Psi_1}, \quad (4.30)$$

and its frequency domain counterpart:

$$\text{FGCI}_{X_2 \rightarrow X_1 | X_3}(f) = \text{FGCI}_{X_2 \Psi_3 \rightarrow \Psi_1}(f). \quad (4.31)$$

To obtain $\text{FGCI}_{X_2 \Psi_3 \rightarrow \Psi_1}(f)$, we need to decompose the variance of $\Psi_1(t)$ into the frequency domain. We compute the Fourier transforms of Eqs. (4.24) and (4.28)

$$\begin{pmatrix} X_1(f) \\ X_3(f) \end{pmatrix} = \overbrace{\begin{pmatrix} G_{11}(f) & G_{13}(f) \\ G_{31}(f) & G_{33}(f) \end{pmatrix}}^{G(f)} \begin{pmatrix} \Psi_1(f) \\ \Psi_3(f) \end{pmatrix}, \quad (4.32)$$

$$\begin{pmatrix} X_1(f) \\ X_2(f) \\ X_3(f) \end{pmatrix} = \begin{pmatrix} H_{11}(f) & H_{12}(f) & H_{13}(f) \\ H_{21}(f) & H_{22}(f) & H_{23}(f) \\ H_{31}(f) & H_{32}(f) & H_{33}(f) \end{pmatrix} \begin{pmatrix} E_1(f) \\ E_2(f) \\ E_3(f) \end{pmatrix}. \quad (4.33)$$

If the spectra $X_1(f)$ and $X_3(f)$ obtained from Eq. (4.32) are identical to the spectra $X_1(f)$ and $X_3(f)$ obtained from Eq. (4.33), then we combine Eq. (4.32) and Eq. (4.33) to yield the following equation:

$$\begin{aligned} \begin{pmatrix} \Psi_1(f) \\ X_2(f) \\ \Psi_3(f) \end{pmatrix} &= \overbrace{\begin{pmatrix} G_{11}(f) & 0 & G_{13}(f) \\ 0 & 1 & 0 \\ G_{31}(f) & 0 & G_{33}(f) \end{pmatrix}}^{G_3^{-1}(f)} \overbrace{\begin{pmatrix} H_{11}(f) & H_{12}(f) & H_{13}(f) \\ H_{21}(f) & H_{22}(f) & H_{23}(f) \\ H_{31}(f) & H_{32}(f) & H_{33}(f) \end{pmatrix}}^{H(f)} \begin{pmatrix} E_1(f) \\ E_2(f) \\ E_3(f) \end{pmatrix} \\ &= \underbrace{\begin{pmatrix} Q_{11}(f) & Q_{12}(f) & Q_{13}(f) \\ Q_{21}(f) & Q_{22}(f) & Q_{23}(f) \\ Q_{31}(f) & Q_{32}(f) & Q_{33}(f) \end{pmatrix}}_{Q(f)} \begin{pmatrix} E_1(f) \\ E_2(f) \\ E_3(f) \end{pmatrix} \end{aligned} \quad (4.34)$$

where $Q(f) = G_3^{-1}(f)H(f)$. After suitable ensemble averaging from Eq. (4.34), the power spectrum of Ψ_1 is found to be:

$$S_{\Psi_1}(f) = Q_{11}(f)\hat{\Sigma}_{11}Q_{11}^*(f) + Q_{12}(f)\hat{\Sigma}_{22}Q_{12}^*(f) + Q_{13}(f)\hat{\Sigma}_{33}Q_{13}^*(f). \quad (4.35)$$

The first term of Eq. (4.35) can be thought as the intrinsic power and the remaining two terms as the combined causal influences of X_2 and Ψ_3 . This interpretation leads immediately to the definition

$$\text{FGCI}_{X_2\Psi_3\rightarrow\Psi_1}(f) = \ln \frac{|S_{\Psi_1}(f)|}{|Q_{11}(f)\hat{\Sigma}_{11}Q_{11}^*(f)|} \quad (4.36)$$

where $S_{\Psi_1}(f)$ is actually the variance of $U_1(t)$ in Eq. (4.23), that is $\Gamma_{X_1|X_1^-X_3^-} = \text{var}(U_1(t))$. Considering the identification in Eq. (4.31), the final expression for the conditional causality from X_2 to X_1 conditional on X_3 is

$$\text{FGCI}_{X_2\rightarrow X_1|X_3}(f) - C = \ln \frac{\Gamma_{X_1|X_1^-X_3^-}}{|Q_{11}(f)\hat{\Sigma}_{11}Q_{11}^*(f)|}. \quad (4.37)$$

The above derivation is made possible by the key assumption that the spectra $X_1(f)$ and $X_3(f)$ coming from Eq. (4.32) and Eq. (4.33) are identical. This certainly holds true on purely theoretical grounds, and it can remain true for simple mathematical systems. For actual physiological signals, however, this condition is hard to satisfy numerically due to practical estimation errors. In order to overcome this problem, a partition matrix technique is presented in [Chen 2006].

For three blocks of time series $x_1(t)$, $x_2(t)$ and $x_3(t)$, we can fit a three-variable VAR model as in Eq. (4.28) and we can also derive its frequency domain expression as in Eq. (4.33).

From Eq. (4.33), writing an expression only for $X_1(f)$ and $X_3(f)$ (making partitions) we have:

$$\begin{pmatrix} X_1(f) \\ X_3(f) \end{pmatrix} = \begin{pmatrix} H_{11}(f) & H_{13}(f) \\ H_{31}(f) & H_{33}(f) \end{pmatrix} \begin{pmatrix} \bar{E}_1(f) \\ \bar{E}_3(f) \end{pmatrix} \quad (4.38)$$

where $\bar{E}_1(f)$ and $\bar{E}_3(f)$ have the following moving average expression:

$$\begin{pmatrix} \bar{E}_1(f) \\ \bar{E}_3(f) \end{pmatrix} = \begin{pmatrix} E_1(f) \\ E_3(f) \end{pmatrix} + \begin{pmatrix} H_{11}(f) & H_{13}(f) \\ H_{31}(f) & H_{33}(f) \end{pmatrix}^{-1} \begin{pmatrix} H_{12}(f) \\ H_{32}(f) \end{pmatrix} E_2(f). \quad (4.39)$$

Given that

$$\begin{pmatrix} \bar{H}_{12}(f) \\ \bar{H}_{32}(f) \end{pmatrix} = \begin{pmatrix} H_{11}(f) & H_{13}(f) \\ H_{31}(f) & H_{33}(f) \end{pmatrix}^{-1} \begin{pmatrix} H_{12}(f) \\ H_{32}(f) \end{pmatrix}, \quad (4.40)$$

we get the covariance matrix of the noise terms given in Eq. (4.39):

$$\bar{\Sigma}(f) = \begin{pmatrix} \Sigma_{11} & \Sigma_{13} \\ \Sigma_{31} & \Sigma_{33} \end{pmatrix} + \begin{pmatrix} \bar{H}_{12}(f) \\ \bar{H}_{32}(f) \end{pmatrix} \begin{pmatrix} \Sigma_{12} & \Sigma_{32} \end{pmatrix} + \begin{pmatrix} \Sigma_{12} \\ \Sigma_{32} \end{pmatrix} \begin{pmatrix} \bar{H}_{12}^*(f) & \bar{H}_{32}^*(f) \end{pmatrix} + \Sigma_{22} \begin{pmatrix} \bar{H}_{12}(f) \\ \bar{H}_{32}(f) \end{pmatrix} \begin{pmatrix} \bar{H}_{12}^*(f) & \bar{H}_{32}^*(f) \end{pmatrix}. \quad (4.41)$$

This covariance matrix is no longer a real matrix, but it is a Hermitian matrix, *i.e.* $\bar{\Sigma}_{13}(f) = \bar{\Sigma}_{31}^*(f)$. Therefore, we can use the following transformation matrix to normalize the bivariate model of Eq. (4.38):

$$\bar{P} = \begin{pmatrix} 1 & 0 \\ -\bar{\Sigma}_{31}(f)/\Sigma_{11} & 1 \end{pmatrix}. \quad (4.42)$$

Therefore, in correspondence with the normalized form in Eq. (4.32), the transfer matrix $G(f)$ is:

$$G(f) = \begin{pmatrix} H_{11}(f) & H_{13}(f) \\ H_{31}(f) & H_{33}(f) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\bar{\Sigma}_{31}(f)/\Sigma_{11} & 1 \end{pmatrix}^{-1}. \quad (4.43)$$

Taking the expansion form of the matrix $G(f)$ to get the matrix $Q(f) = G_3^{-1}(f)H(f)$, and considering $\Gamma_{X_1|X_1^-X_3^-} = \bar{\Sigma}_{11}$, we can still use Eq. (4.37) to get the conditional causality.

4.2. Phase slope index and associated causality indices

The Phase Slope Index (PSI) measure was originally proposed by Nolte [Nolte 2008] to detect information flow in unidirectional propagation graphs and applied in neuroscience [Cooray 2011, Haufe 2010, Jung 2012, Nolte 2010, Rana 2012, Sekihara 2010]. The idea behind this measure is that the slope of the cross-spectrum phase between two different source activities depends on the time needed for the information flow between those areas and on the corresponding direction. This method estimates the direction of the flow by computing the slope of the phase of the ordinary coherence function. Now, when two time series display direct or indirect relations (Fig. 4.1 (a) and (b)), PSI based on the Ordinary Coherence (OC) function fails to distinguish between these two types of relations. As a matter of fact, when a third channel accounts for the linear relation between two other signals under scrutiny, the amplitude of the coherence function between these two signals is one (as it is when there is a direct linear relation between these signals). To deal with this issue, we propose to replace OC with partial coherence (PC) [Yang 2010]. However, just like ordinary coherence, since the partial coherence function between two signals only carries a single direction's information (given by the phase itself or its opposite value), it leads to a symmetric index and, consequently, is unable to detect bidirectional (feedback) flows. Hence, to mitigate the two previous issues, we investigate new Causality Indices (CI) to detect and differentiate unidirectional and bidirectional relations between multivariate time series: the first one is a new CI based on the Directed COherence (DCOH) function [Saito 1981] when considering pairwise analysis (*i.e.* only two observations are considered at the same time) and the second one is based on the Directed Transfer Function (DTF) [Kamiński 1991] when considering multivariate analysis (*i.e.* joint analysis of more than two signals) [Yang 2011b]. Then, to meet potential direct and indirect relations in bidirectional situations, we recommend to introduce Partial Directed Coherence (PDC) [Baccalá 2001] in a new index [Yang 2012]. Until now, only the amplitudes of these different transfer functions have been considered in the literature to estimate the brain connectivity between structures in the frequency domain. Such approaches obviously failed in differentiating even quite simple scenarios, *e.g.* when two investigated observations

only consisted of different time shifted versions of a third observation. In the subsequent sections, the aforementioned causality indices are detailed theoretically. Note that, in contrast with Granger causality based indices, those based on phase slope index merge causality and time delay. This is an interesting property even if slightly awkward to interpret.

4.2.1. Phase slope index

The PSI basic hypothesis relies on the exploitation of the phase monotony between signals which appears when the frequency components of one signal precede temporally those of another signal [Nolte 2008]. PSI is defined in order to summarize information on the slope of the phase of the cross-spectrum between two processes X_m and X_n . The theoretical idea of this index is to exploit properly relative time delays between spectral components of the two signals X_m and X_n only in the frequency bands where the coherence is significant. Let us consider the functional

$$\Delta_{m,n} = \int \frac{\partial \theta(f)}{\partial f} |C_{mn}(f)|^2 df \quad (4.44)$$

where

$$C_{mn}(f) = S_{mn}(f) / \sqrt{S_{mm}(f)S_{nn}(f)} \quad (4.45)$$

is the OC function between signals X_m and X_n , with $C_{mn}(f) = |C_{mn}(f)| \exp(i\theta(f))$. Clearly

$$|\Delta_{m,n}| \leq \left| \int \frac{\partial \theta(f)}{\partial f} df \right| \text{ as } |C_{mn}(f)|^2 \leq 1 \text{ with strict equality if and only if } |C_{mn}(f)|^2 = 1, \forall f \in [0,1].$$

If this condition is verified, $\Delta_{m,n}$ can be interpreted as a mean phase slope, *i.e.* as a mean time delay. $\Delta_{m,n}$ increases when the mean time delay increases and also when the modulus of the coherence function increases, and reciprocally. It represents a compromise between acquisition of information on time delay and propagation direction discarding frequencies that could lead to erroneous phase estimations.

It is easy to verify, as shown below, that, for small values of δf , a possible numerical approximation of $\Delta_{m,n}$ is

$$\Delta_{m,n} \approx \text{Im} \left(\sum_{f \in F_d} C_{mn}^*(f) C_{mn}(f + \delta f) \right) \quad (4.46)$$

which corresponds to the Phase Slope Index defined in [Nolte 2008] and noted PSI_{mn} hereafter.

$\text{Im}(\cdot)$ denotes the imaginary part and $*$ represents the conjugate value; F_d is a discrete set of frequencies over which the index is computed and which can be chosen by the experimenter according to some knowledge on the signals' characteristics. For example, if it is known that the signals are band limited, F_d can be reduced to only some critical frequencies. Similarly, when using FFT (Fast Fourier Transform) and without any *a priori* knowledge on the signals, the maximum set in

the normalized half frequency band is given by: $F_d = \{0, 1/N, \dots, 1/2 - \delta f\}$ where $\delta f = 1/N$ (in this case the frequential step-size δf corresponds to the frequency resolution $1/N$). The auto-spectral density functions $S_{mm}(f)$ and $S_{nn}(f)$ are the Fourier transforms of the auto-correlation of the signals $X_m(t)$ and $X_n(t)$ respectively, the cross-spectral density function $S_{mn}(f)$ is the Fourier transform of the cross-correlation between signals $X_m(t)$ and $X_n(t)$ defined by

$$\rho_{mn} = E[X_m(t)X_n^*(t - \tau)] \quad (4.47)$$

where $E[\bullet]$ is the expectation operator and τ is a time displacement. With

$$C_{mn}(f) = |C_{mn}(f)| \exp(i\theta(f)), \quad (4.48)$$

we can write

$$\text{PSI}_{mn} = \sum_{f \in F} \sin(\theta(f + \delta f) - \theta(f)) |C_{mn}(f)C_{mn}(f + \delta f)|, \quad (4.49)$$

which can be approximated, for a sufficiently small δf , by

$$\text{PSI}_{mn} \approx \sum_{f_i \in F_d} (\theta(f_i + \delta f) - \theta(f_i)) |C_{mn}(f_i)|^2 \quad (4.50)$$

or

$$\text{PSI}_{mn} \approx \int_{f \in F} \frac{\partial \theta(f)}{\partial f} |C_{mn}(f)|^2 df \quad (4.51)$$

where the second term in Eq. (4.50) corresponds to a Riemann sum which approximates the continuous sum in Eq. (4.51) on a continuous range F . Clearly, the magnitudes of the coherence function provide for weighting the phase difference between two consecutive frequencies and, consequently, decrease its impact when one of them (or both) is (are) low. The sign of PSI indicates the flow direction and its magnitude increases along with the delay and the coherence module. Clearly, this index (i) only works in situations of unidirectional connections, and (ii) cannot discriminate between direct and indirect relations. Following our notations, a positive value of PSI_{mn} means that the signal X_n is a delayed version of X_m . As it is well known, the linearity of the phase corresponds to a pure delay between signals X_m and X_n . When one signal contributes to the second with multiple, different delays, the phase becomes nonlinear (the slope is no longer a constant).

In the following, PSI using the OC function is named PSI-OC and given by:

$$\text{PSI}_{mn\text{-OC}} = \text{Im} \left(\sum_{f \in F} \text{OC}_{mn}^*(f) \text{OC}_{mn}(f + \delta f) \right). \quad (4.52)$$

Given Eq. (4.47), the auto-spectral and cross-spectral density functions may be obtained by two different techniques, either from direct Fourier transforms of the observed realizations $x_m(t)$ and $x_n(t)$, as in [Nolte 2008], or from AR modeling. In the first one, the expectation required to get the spectral density functions is obtained by averaging and overlap. In the second one, the methodology for the estimation in the multivariate case can be derived as follows.

Let X_1, X_2, \dots, X_Q be Q zero-mean signals whose discrete-time observations are noted $x_1(t), x_2(t), \dots, x_Q(t)$, $t = 1, 2, \dots, T$, where T is the signal length. We can model these observations by a Q dimensional multivariate AR model of order p (same form as Eq. (4.7)):

$$\begin{pmatrix} x_1(t) \\ \vdots \\ x_Q(t) \end{pmatrix} = \sum_{k=1}^p \begin{pmatrix} \hat{\alpha}_{11}(k) & \hat{\alpha}_{12}(k) & \cdots & \cdots & \hat{\alpha}_{1Q}(k) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \hat{\alpha}_{mn}(k) & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \hat{\alpha}_{Q1}(k) & \cdots & \cdots & \cdots & \hat{\alpha}_{QQ}(k) \end{pmatrix} \begin{pmatrix} x_1(t-k) \\ \vdots \\ x_Q(t-k) \end{pmatrix} + \begin{pmatrix} \hat{w}_1(t) \\ \vdots \\ \hat{w}_Q(t) \end{pmatrix}. \quad (4.53)$$

Using the lag operator L , $L^k x_m(t) = x_m(t-k)$, with $m = 1, 2, \dots, Q$, and $k = 1, 2, \dots, p$, we have

$$\begin{pmatrix} 1 - \sum_{k=1}^p \hat{\alpha}_{11}(k)L^k & -\sum_{k=1}^p \hat{\alpha}_{12}(k)L^k & \cdots & -\sum_{k=1}^p \hat{\alpha}_{1Q}(k)L^k \\ -\sum_{k=1}^p \hat{\alpha}_{21}(k)L^k & 1 - \sum_{k=1}^p \hat{\alpha}_{22}(k)L^k & \cdots & -\sum_{k=1}^p \hat{\alpha}_{2Q}(k)L^k \\ \vdots & \vdots & \ddots & \vdots \\ -\sum_{k=1}^p \hat{\alpha}_{Q1}(k)L^k & -\sum_{k=1}^p \hat{\alpha}_{Q2}(k)L^k & \cdots & 1 - \sum_{k=1}^p \hat{\alpha}_{QQ}(k)L^k \end{pmatrix} \begin{pmatrix} x_1(t) \\ \vdots \\ x_Q(t) \end{pmatrix} = \begin{pmatrix} \hat{w}_1(t) \\ \vdots \\ \hat{w}_Q(t) \end{pmatrix}. \quad (4.54)$$

Applying Fourier transform to both sides of Eq. (4.54) leads to

$$\underbrace{\begin{pmatrix} D_{11}(f) & D_{12}(f) & \cdots & D_{1Q}(f) \\ D_{21}(f) & D_{22}(f) & \cdots & D_{2Q}(f) \\ \vdots & \vdots & \ddots & \vdots \\ D_{Q1}(f) & D_{Q2}(f) & \cdots & D_{QQ}(f) \end{pmatrix}}_{D(f)} \begin{pmatrix} X_1(f) \\ \vdots \\ X_Q(f) \end{pmatrix} = \begin{pmatrix} \hat{W}_1(f) \\ \vdots \\ \hat{W}_Q(f) \end{pmatrix} \quad (4.55)$$

where $\hat{W}(f)$ denotes estimated frequency component (keeping the same convention as explained in section 4.1.2). The components of the coefficient matrix $D(f)$ are

$$D_{mn}(f) = \begin{cases} 1 - \sum_{k=1}^p \hat{\alpha}_{mn}(k) e^{-2i\pi f k}, & m = n \\ -\sum_{k=1}^p \hat{\alpha}_{mn}(k) e^{-2i\pi f k}, & m \neq n \end{cases}. \quad (4.56)$$

Defining the transfer function $H(f)$ as the inverse of the coefficient matrix $D(f)$, we obtain

$$\begin{pmatrix} X_1(f) \\ \vdots \\ X_Q(f) \end{pmatrix} = \underbrace{\begin{pmatrix} H_{11}(f) & H_{12}(f) & \cdots & H_{1Q}(f) \\ H_{21}(f) & H_{22}(f) & \cdots & H_{2Q}(f) \\ \vdots & \vdots & \ddots & \vdots \\ H_{Q1}(f) & H_{Q2}(f) & \cdots & H_{QQ}(f) \end{pmatrix}}_{H(f)} \begin{pmatrix} \hat{W}_1(f) \\ \vdots \\ \hat{W}_Q(f) \end{pmatrix}. \quad (4.57)$$

Then, we get the spectral density matrix

$$S(f) = \begin{pmatrix} S_{11}(f) & S_{12}(f) & \cdots & S_{1Q}(f) \\ S_{21}(f) & S_{22}(f) & \cdots & S_{2Q}(f) \\ \vdots & \vdots & \ddots & \vdots \\ S_{Q1}(f) & S_{Q2}(f) & \cdots & S_{QQ}(f) \end{pmatrix} = H(f) \Gamma H^\dagger(f) \quad (4.58)$$

where $\Gamma = \begin{pmatrix} \text{var}(\hat{W}_1(t)) & \text{cov}(\hat{W}_1(t), \hat{W}_2(t)) & \cdots & \text{cov}(\hat{W}_1(t), \hat{W}_Q(t)) \\ \text{cov}(\hat{W}_2(t), \hat{W}_1(t)) & \text{var}(\hat{W}_2(t)) & \cdots & \text{cov}(\hat{W}_2(t), \hat{W}_Q(t)) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(\hat{W}_Q(t), \hat{W}_1(t)) & \text{cov}(\hat{W}_Q(t), \hat{W}_2(t)) & \cdots & \text{var}(\hat{W}_Q(t)) \end{pmatrix}$, $\text{var}(\cdot)$ (resp. $\text{cov}(\cdot)$)

here stands for estimated variance (resp. covariance) [Ding 2006], and the symbol \dagger denotes Hermitian transpose. Finally, the corresponding PSI-OC can be calculated using Eqs. (4.58), (4.45) and (4.52).

To deal with a more general situation, we extend this idea to other coupling based functions as the phase is monotonic to propose novel causality indices, noted CI afterwards.

4.2.2. Causality index using partial coherence

The partial coherence function gives the level of coupling between two signals X_m and X_n when the influence of the $Q - 2$ other signals is removed [Bendat 1986a]. It is defined by

$$PC_{mn \cdot X_{Q \setminus 2}}(f) \triangleq \frac{S_{mn \cdot X_{Q \setminus 2}}(f)}{\sqrt{S_{mm \cdot X_{Q \setminus 2}}(f) S_{nn \cdot X_{Q \setminus 2}}(f)}} \quad (4.59)$$

where $X_{Q \setminus 2} \triangleq X_1 \cdots X_{m-1} X_{m+1} \cdots X_{n-1} X_{n+1} \cdots X_Q$. $S_{mn \cdot X_{Q \setminus 2}}(f)$ is the conditional cross-spectral density function between signals X_m and X_n given $X_{Q \setminus 2}$. $S_{mm \cdot X_{Q \setminus 2}}(f)$ and $S_{nn \cdot X_{Q \setminus 2}}(f)$ are conditional auto-spectral density functions of signals X_m and X_n respectively. In [Bendat 1986a], Bendat and Piersol gave the methodology to obtain the expressions of $S_{mn \cdot X_{Q \setminus 2}}(f)$, $S_{mm \cdot X_{Q \setminus 2}}(f)$, and $S_{nn \cdot X_{Q \setminus 2}}(f)$. This methodology is introduced simply in the following for $m = 1$ and $n = 2$. Rewriting Eq. (4.58) in the following form

$$\left(\begin{array}{cc|cccc} S_{11}(f) & S_{12}(f) & S_{13}(f) & S_{14}(f) & \cdots & S_{1Q}(f) \\ S_{21}(f) & S_{22}(f) & S_{23}(f) & S_{24}(f) & \cdots & S_{2Q}(f) \\ \hline S_{31}(f) & S_{32}(f) & S_{33}(f) & S_{34}(f) & \cdots & S_{3Q}(f) \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ S_{Q1}(f) & S_{Q2}(f) & S_{Q3}(f) & S_{Q4}(f) & \cdots & S_{QQ}(f) \end{array} \right) = \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}, \quad (4.60)$$

the conditional spectral matrix between signals $x_1(t)$ and $x_2(t)$ given X_{Q-2} is defined by

$$\begin{pmatrix} S_{11 \cdot X_{Q/2}}(f) & S_{12 \cdot X_{Q/2}}(f) \\ S_{21 \cdot X_{Q/2}}(f) & S_{22 \cdot X_{Q/2}}(f) \end{pmatrix} = B_1 - B_2 B_4^{-1} B_3. \quad (4.61)$$

To obtain a conditioned version of PSI defined in Eq. (4.52), we can now replace the coherence function with the partial coherence and define a causality index, named CI-PC: the influence of the $Q-2$ other signals is removed and only the direct influence between X_m and X_n is considered:

$$CI_{mn}\text{-PC} = \text{Im} \left(\sum_{f \in F} PC_{mn \cdot X_{Q/2}}^*(f) PC_{mn \cdot X_{Q/2}}(f + \delta f) \right). \quad (4.62)$$

4.2.3. Causality index using directed coherence and directed transfer function

While ordinary and partial coherences focus on mutual interaction between structures, directed coherence, as well as directed transfer function, is referring to the concept of Granger causality. Consequently, unlike ordinary and partial coherences, directed coherence and directed transfer function are asymmetric quantities. The concept of directed coherence was first developed by Saito and Harashima [Saito 1981] to jointly analyze information production in two time series, each having its proper white noise source, which can be seen as a local innovation, as well as a common source, seen as an external innovation. While coherence measures the degree of linear correlation as a total, the "directed coherences" can be seen as "correlations with direction" between the two observed signals expressed in the frequency domain and, interestingly, be regarded as two contributing weighted factors in the expression of the global coherence. Given these two observations, viewed as the outputs of the system, these coherences describe the connection between the first (resp. second) noise input and the second (resp. first) output of this system.

To formulate this problem of direction of correlating influences, Saito and Harashima [Saito 1981] considered a bivariate autoregressive process (composed of two realizations $x_1(t)$ and $x_2(t)$) of order p including a common noise source $w_3(t)$ such as:

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \sum_{k=1}^p \begin{pmatrix} \alpha_{11}(k) & \alpha_{12}(k) \\ \alpha_{21}(k) & \alpha_{22}(k) \end{pmatrix} \begin{pmatrix} x_1(t-k) \\ x_2(t-k) \end{pmatrix} + \begin{pmatrix} b_{11} & b_{13} & 0 \\ 0 & b_{23} & b_{22} \end{pmatrix} \begin{pmatrix} w_1(t) \\ w_3(t) \\ w_2(t) \end{pmatrix}. \quad (4.63)$$

where b_{mn} , $m \in \{1,2\}$, $n \in \{1,2,3\}$, are weight factors, and $w_j(t)$, $j=1,2,3$ are realizations of independent zero mean white Gaussian noises of unit variance.

In the frequency domain, we have:

$$\begin{aligned}
\begin{pmatrix} X_1(f) \\ X_2(f) \end{pmatrix} &= \begin{pmatrix} 1 - \sum_{k=1}^p \alpha_{11}(k) e^{-2i\pi f k} & -\sum_{k=1}^p \alpha_{12}(k) e^{-2i\pi f k} \\ -\sum_{k=1}^p \alpha_{21}(k) e^{-2i\pi f k} & 1 - \sum_{k=1}^p \alpha_{22}(k) e^{-2i\pi f k} \end{pmatrix}^{-1} \begin{pmatrix} b_{11} & b_{13} & 0 \\ 0 & b_{23} & b_{22} \end{pmatrix} \begin{pmatrix} W_1(f) \\ W_3(f) \\ W_2(f) \end{pmatrix} \\
&= \begin{pmatrix} A_{11}(f) & A_{12}(f) \\ A_{21}(f) & A_{22}(f) \end{pmatrix} \begin{pmatrix} b_{11} & b_{13} & 0 \\ 0 & b_{23} & b_{22} \end{pmatrix} \begin{pmatrix} W_1(f) \\ W_3(f) \\ W_2(f) \end{pmatrix}.
\end{aligned} \tag{4.64}$$

Then, we introduce the matrix $H(f)$ such as:

$$H(f) = \begin{pmatrix} H_{11}(f) & H_{13}(f) & H_{12}(f) \\ H_{21}(f) & H_{23}(f) & H_{22}(f) \end{pmatrix} = \begin{pmatrix} A_{11}(f) & A_{12}(f) \\ A_{21}(f) & A_{22}(f) \end{pmatrix} \begin{pmatrix} b_{11} & b_{13} & 0 \\ 0 & b_{23} & b_{22} \end{pmatrix} \tag{4.65}$$

where $H_{mn}(f)$, $m \in \{1,2\}$, $n \in \{1,2,3\}$ are the transfer functions of the system. The DCOH estimate of the linear feedback from the innovation process $w_m(t)$ corresponding to the observation $x_m(t)$ to the observed signal $x_n(t)$, with $m \neq n$, m and n in $\{1,2\}$, is defined by

$$DCOH_{mn}(f) \triangleq \frac{H_{nm}(f)}{\sqrt{\sum_{j=1,2,3} |H_{nj}(f)|^2}}. \tag{4.66}$$

Following the previous idea on PSI, we define a causality index, named CI-DCOH, as follows:

$$CI_{mn}\text{-DCOH} = \text{Im} \left(\sum_{f \in F} DCOH_{mn}^*(f) DCOH_{mn}(f + \delta f) \right). \tag{4.67}$$

Contrary to the PSI-OC, this index is asymmetric allowing the detection of bidirectional flows. It is no longer relative to the slope of the phase between the observations themselves but to those between the noise sources W_m and the signals X_n , with $m \neq n \in \{1,2\}$.

Following the concept of directed coherence, the directed transfer function was introduced by Kamiński and Blinowska [Kamiński 1991] to deal with a number of observations greater than two. In this case, contrary to the previous situation, no hidden common noise source is considered, and each observation can be viewed as produced by its own innovation sequence linearly combined with delayed versions of all observations. In the same manner as previously, directed transfer function from the i -th input to the j -th output of the system can be derived. Let us indicate that both estimators are normalized with respect to the structure that receives the signal. Therefore, in the framework of multivariate observations, we extend the aforementioned concept of causality index to Directed Transfer Function (DTF) (instead of DCOH).

From Eq. (4.57), the DTF from channel m to channel n is defined by

$$DTF_{mn}(f) \triangleq \frac{H_{nm}(f)}{\sqrt{\sum_{m=1}^Q |H_{nm}(f)|^2}} \tag{4.68}$$

where $H_{nm}(f)$ is the (n,m) element of the matrix $H(f)$, corresponding to the normalized contribution of the input sequence w_m onto the output signal x_n . DTF_{mn} is a multichannel causality measure based on the AR model and measures the flow from channel m to channel n . It is constructed using elements of the transfer matrix of the AR model. This matrix is not symmetrical and, the value of the DTF, at a given frequency, represents a ratio between the inflow to channel n from channel m to all the inflows to channel n . So, contrary to ordinary coherence, it can deal with bidirectional flow and answer the question of effective connectivity. DTF itself is able to quantify the strength of such connectivity but the supplementary interest of the PSI based indicator relies on the phase information carried by each element of the transfer function used in the DTF. The relevance of causality index using DTF consists in using on the one hand the amplitude of the directed transfer function and on the other hand the variation of its phase between two adjacent frequencies. So, averaging over the frequency band allows assessing not only the strength of the coupling but also the importance of the delay. In the same way as previously, we define a CI based on DTF, noted CI-DTF, as follows:

$$CI_{mn}\text{-DTF} = \text{Im} \left(\sum_{f \in F} DTF_{mn}^*(f) DTF_{mn}(f + \delta f) \right). \quad (4.69)$$

4.2.4. Causality index using partial directed coherence

Following the above developments, in [Baccalá 2001], Baccalá and Sameshima contrasted partial directed coherence with directed coherence to show how partial directed coherence provides direct structural information for multivariate time signals, as partial coherence does compared to ordinary coherence in unidirectional flow models. So, given $Q > 2$ observations, the partial directed coherence function describes the interaction between two of these observations when the influence due to all other $Q - 2$ time series is discounted.

According to Eq. (4.55), the PDC from m to n conditionally to other $Q - 2$ signals can be written

$$PDC_{mn}(f) \triangleq \frac{D_{nm}(f)}{\sqrt{d_m^\dagger(f) \cdot d_m(f)}} \quad (4.70)$$

where $D_{nm}(f)$ is the (n,m) element of $D(f)$ and $d_m(f)$ is the m th column of $D(f)$. $D_{nm}(f)$ represents the contribution of the past of the signal X_m on the present signal X_n . The sign \dagger denotes conjugate transpose. The arguments to define a PDC based causality index can be fully justified as we want to develop an average measure (i) to quantify the relative delays between multiple signals and not only to qualify them as the PDC magnitude did, and (ii) to weight properly different frequency regions according to the strength of the direct coupling. Consequently, we propose to define a PDC based CI measure, noted CI-PDC, by

$$CI_{mn}\text{-PDC} = \text{Im} \left(\sum_{f \in F} PDC_{mn}^*(f) PDC_{mn}(f + \delta f) \right). \quad (4.71)$$

If the only contribution of the signal X_m into the signal X_n is due to a given delay, the phase of $D_{nm}(f)$ as well as the phase of PDC is linear, so that CI-PDC represents the slope of this phase, weighted by the magnitude of PDC.

4.3. Transfer entropy and conditional transfer entropy

A theoretical information measure, namely Transfer Entropy (TE), was originally proposed by Schreiber [Schreiber 2000] to identify the direction of the information flow and to quantify the strength of coupling between complex systems. This "model-free" technique is based on the transition probabilities between states of the considered systems, where states are defined from observed signals. During the last decennium, this method had been applied and tested in some nonlinear benchmark models, real EEG signals and magnetoencephalographic data in cortical and cerebellar networks with no ground reference [Sabesan 2009b, Vicente 2011, Wibral 2011a]. Moreover, from these foregoing literatures, it appears that certain calibration parameters involved in the TE estimation play a significant role in acquiring the correct information flow direction between two systems (driven system and driving system), for instance, the calibration parameters of two Markov process orders for the driven system and for the driving one. In [Schreiber 2000], only the Markov process order of the driving system is mentioned: it is fixed to that of the driven one or set to 1 with the reason that the latter is preferable for computational reasons. In [Sabesan 2007], Sabesan *et al.* further investigated this problem. Two measures, the delayed Mutual Information (MI) and Autocorrelation Function (AF) [Fraser 1986, Martinerie 1992], were proposed to estimate the Markov process order of the driven system (this approach can be discussed because, for a Markov process, the correlation time can be different from the model order). The Markov process order of the driving system was set to 1 since they presumed that the current state of the driving system is sufficient to arouse a considerable change in the dynamics of the driven system with one time step in the future. Later, Vicente *et al.* [Vicente 2011] and Wibral *et al.* [Wibral 2011a] used Cao's criterion [Cao 1997] based on the false neighbour computation and autocorrelation function to estimate the Markov process order of the driven system for TE estimation (but they did not mention the estimation of the order for the driving system). In 1981, Katz [Katz 1981] also suggested to use standard linear measures for AR model order estimation – AIC and BIC – to estimate the order of Markov processes. However, AIC and BIC are generally used to select one single value for the Markov process orders of driven and driving systems. In most situations, the Markov process orders of these two systems are different (this is typically the case when there is no influence of one system on another). Consequently, it would lead to some erroneous estimation in TE when the Markov process orders of the driven and driving systems are chosen identical [Yang 2011a]. To deal with this issue, our first improvement in the TE measure consists in preferring two criteria, namely the generalized AIC (gAIC) and the generalized BIC (gBIC) respectively, to estimate the Markov process order so as to enhance the correctness and robustness of TE estimation. As we discussed in the aforementioned sections, similar to GCI, TE is a method to characterize causal interactions by pairwise analysis. Therefore, it cannot differentiate direct or indirect relation in multivariable complex systems. To solve this problem, we consider another technique, called Conditional Transfer Entropy (CTE), to discriminate direct and indirect relations.

Hereafter, the methodology of TE is first recalled in section 4.3.1, and the methodology of two criteria – gAIC and gBIC – is developed to determine how to estimate the different orders for driven and driving systems. Finally, the CTE approach is detailed in section 4.3.2.

4.3.1. *Transfer entropy*

Let us briefly recall the most basic concepts of information theory [Billingsley 1965]. The average number of bits needed to optimally encode independent draws of the discrete state variable X following a probability distribution $P(x)$ is given by the Shannon entropy [Shannon 1949] $H_X = -\sum_x P(x) \log_2 P(x)$, where the sum extends over all states x the process can assume. The base of the logarithm determines only the units used for measuring information and will be dropped henceforth^{IV}.

In order to construct an optimal encoding that uses just as many bits as given by the entropy, it is necessary to know the probability distribution $P(x)$. The excess number of bits that will be coded if a different distribution $Q(x)$ is used is given by the Kullback entropy [Kullback 1997] $K_X = \sum_x P(x) \ln(P(x)/Q(x))$. We will later also need the Kullback entropy for conditional probabilities $p(x|y)$. For a given state value y we have $K_Y = \sum_x P(x|y) \ln(P(x|y)/Q(x|y))$. Summation over y with respect to $P(y)$ yields to the conditional version of the Kullback entropy:

$$K_{X|Y} = \sum_{x,y} P(x,y) \ln \frac{P(x|y)}{Q(x|y)}. \quad (4.72)$$

The mutual information of two processes X and Y with joint probability $P_{XY}(x,y)$ can be seen as the excess amount of code produced by erroneously assuming that the two systems are independent, *i.e.* using $Q_{XY}(x,y) = P_X(x)P_Y(y)$ instead of $P_{XY}(x,y)$. The corresponding Kullback entropy is

$$MI_{XY} = \sum_{x,y} P(x,y) \ln \frac{P(x,y)}{P(x)P(y)}, \quad (4.73)$$

which is the well known formula for the mutual information. Note that MI is symmetric under the exchange of X and Y and therefore does not contain any directional sense.

A related, non symmetric quantity is the conditional entropy $H_{X|Y} = -\sum P(x,y) \ln P(x|y) = H_{XY} - H_Y$. However, since $H_{X|Y} - H_{Y|X} = H_X - H_Y$, it is non symmetric only due to the different individual entropies and not due to information flow. By denoting x_n and y_n two time samples from signals X and Y , mutual information can be given a time directional sense in a somewhat *ad hoc* way by introducing a time lag in either one of the variables and compute, *e.g.*,

^{IV} In this section, for reason of clarity, we use variables X , Y and Z instead of X_1 , X_2 and X_3 as in the previous sections.

$$MI_{XY}(\tau) = \sum_{x_n, y_{n-\tau}} P(x_n, y_{n-\tau}) \ln \frac{P(x_n, y_{n-\tau})}{P(x)P(y)}. \quad (4.74)$$

As we will see below, considering the two systems at different times occurs naturally as soon as transition probabilities are introduced. One can specify the dynamical structure of a signal by introducing transition probabilities rather than static probabilities. If we consider a signal X such that the conditional probability to find X in state x_{n+1} at time $n+1$ is independent of the states $x_{n-p}, p \geq k$: $P(x_{n+1} | x_n, \dots, x_{n-k+1}) = P(x_{n+1} | x_n, \dots, x_{n-k+1}, x_{n-k}, \dots)$, X is called a (discrete time) Markov process. Moreover, if $P(x_{n+1} | x_n, \dots, x_{n-k+1})$ is invariant with respect to n , X is a stationary Markov process.

Henceforth, we use the shorthand notation $x_n^{(k)} = (x_n, \dots, x_{n-k+1})$. The average number of bits needed to encode one additional state of the system if all previous states are known is given by the entropy rate

$$h_X = - \sum_{x_{n+1}, x_n^{(k)}} P(x_{n+1}, x_n^{(k)}) \ln P(x_{n+1} | x_n^{(k)}). \quad (4.75)$$

Since $P(x_{n+1} | x_n^{(k)}) = P(x_{n+1}^{(k+1)}) / P(x_n^{(k)})$, this is just the difference between the Shannon entropies of the processes given by $k+1$ and k dimensional delay vectors [Kantz 1997] constructed from X : $h_X = H_{X_n^{(k+1)}} - H_{X_n^{(k)}}$.

For the study of the dynamics of shared information between processes it is desirable to generalize the entropy rate, rather than Shannon entropy, to more than one system, since the dynamics of the processes is contained in the transition probabilities. The most straightforward way to construct a mutual information rate by generalizing h_X to two processes (X, Y) is again by measuring the deviation from independence. The corresponding Kullback entropy is still symmetric under the exchange of X and Y . Therefore, in the following, after briefly summarizing the above content, the methodology of transfer entropy is developed.

Hereafter, let us use more general mathematical notations to encompass all types of probability distribution (e.g. continuous, discrete). Considering a k -th order Markov process X , we have

$$\forall k' > k: P_{X_{n+1}/X_n^{(k)}=x_n^{(k)}}(dx_{n+1}) = P_{X_{n+1}/X_n^{(k')}=x_n^{(k')}}(dx_{n+1}). \quad (4.76)$$

By considering the auxiliary random process Y , the relation in Eq. (4.76) can be extended to formalize the absence of information flow from Y to X (for given k and l values):

$$P_{X_{n+1}/X_n^{(k)}=x_n^{(k)}}(\cdot) = P_{X_{n+1}/X_n^{(k)}=x_n^{(k)}, Y_n^{(l)}=y_n^{(l)}}(\cdot). \quad (4.77)$$

The deviation from this assumption can be quantified using the Kullback pseudo-metric, which leads to define the transfer entropy from Y to X :

$$TE_{Y \rightarrow X} = \int_{\mathbb{R}^{k+l+1}} P_{x_{n+1}, x_n^{(k)}, y_n^{(l)}} \left(dx_{n+1}, dx_n^{(k)}, dy_n^{(l)} \right) \times \log_2 \left(\left[\frac{dP_{x_{n+1}/x_n^{(k)}, y_n^{(l)}}}{dP_{x_{n+1}/x_n^{(k)}}} \right] (x_{n+1}) \right) \quad (4.78)$$

where the ratio in Eq. (4.78) corresponds to a derivative of the numerator conditional measure with respect to the denominator one. This measure is not symmetric. TE can be estimated from observations (x_n, y_n) , $n = 1, 2, \dots, N$, using a kernel discrete estimation of $(x_{n+1}, x_n^{(k)}, y_n^{(l)})$ distribution [Schreiber 2000]:

$$\hat{P}(x_{n+1}, x_n^{(k)}, y_n^{(l)}) \propto \sum_m \Theta \left(\left\| \begin{pmatrix} x_{n+1} - x_{m+1} \\ x_n^{(k)} - x_m^{(k)} \\ y_n^{(l)} - y_m^{(l)} \end{pmatrix} \right\| - r \right) \triangleq C_{n,r}, \quad n = 1, 2, \dots, N-1 \quad (4.79)$$

which depends on a neighborhood size (radius r). Then, it can be used to compute the estimation

$$\widehat{TE}_{Y \rightarrow X} = \sum_n \hat{P}(x_{n+1}, x_n^{(k)}, y_n^{(l)}) \log_2 \frac{\hat{P}(x_{n+1} | x_n^{(k)}, y_n^{(l)})}{\hat{P}(x_{n+1} | x_n^{(k)})} \quad (4.80)$$

where conditional probabilities in Eq. (4.80) are obtained from estimated joint probabilities in Eq. (4.79). $\Theta(\cdot)$ is defined by $\Theta(x > 0) = 1$, $\Theta(x \leq 0) = 0$; $\|\bullet\|$ is the sup-norm and summation is performed on $[1, 2, \dots, N-1]$.

Because of Bayes' rule $P(A|B) = P(A, B) / P(B)$, the transfer entropy estimation in Eq. (4.80) can be rewritten as

$$\widehat{TE}_{Y \rightarrow X} = \sum_n \hat{P}(x_{n+1}, x_n^{(k)}, y_n^{(l)}) \log_2 \frac{\hat{P}(x_{n+1}, x_n^{(k)}, y_n^{(l)}) \hat{P}(x_n^{(k)})}{\hat{P}(x_n^{(k)}, y_n^{(l)}) \hat{P}(x_{n+1}, x_n^{(k)})}. \quad (4.81)$$

Considering the more complicated issue of real data [Schnupp 2006], we define a selective index of the flow direction, named PTE, by

$$PTE_{Y \rightarrow X} = \frac{|\widehat{TE}_{Y \rightarrow X}|}{|\widehat{TE}_{X \rightarrow Y}| + |\widehat{TE}_{Y \rightarrow X}|}, \quad (4.82)$$

$$\widehat{PTE}_{X \rightarrow Y} = \frac{|\widehat{TE}_{X \rightarrow Y}|}{|\widehat{TE}_{X \rightarrow Y}| + |\widehat{TE}_{Y \rightarrow X}|} \quad (4.83)$$

to help in quantifying the contribution of a signal to another one. When $\widehat{PTE}_{Y \rightarrow X}$ (resp. $\widehat{PTE}_{X \rightarrow Y}$) equals 1, only Y drives (causes) X (resp. X drives Y). Values of $\widehat{PTE}_{Y \rightarrow X}$ and $\widehat{PTE}_{X \rightarrow Y}$ equal to 0.5 indicate that there exists an equivalent bidirectional flow.

4.3.1.1. Selection of k and l

The values of the parameters k and l are the orders of the Markov process for the two coupled processes X and Y [Sabesan 2009b]. The selection of these parameters plays an important role to expect dependable values for transfer entropy. In the standard definition of AIC and BIC, the orders k and l are identical, which can not only produce extra computation but also result in some erroneous estimation in TE. Facing this problem, we prefer the two measures, the generalized Akaike Information Criterion (gAIC) and the generalized Bayesian Information Criterion (gBIC) to choose independently different values for the parameters k and l . The procedures of gAIC and gBIC are detailed hereafter:

Rewriting Eq. (D.1) with a 2-dimensional VAR process $[z_n(1), z_n(2)]^T = [x_n, y_n]^T$, we have

$$\begin{cases} x_n = \sum_{i=1}^{q_{xx}} \alpha_{xx}(i) x_{n-i} + \sum_{i=1}^{q_{yx}} \alpha_{yx}(i) y_{n-i} + w_{n,x} \\ y_n = \sum_{i=1}^{q_{xy}} \alpha_{xy}(i) x_{n-i} + \sum_{i=1}^{q_{yy}} \alpha_{yy}(i) y_{n-i} + w_{n,y} \end{cases} \quad (4.84)$$

From the definitions of standard AIC (Eq. (D.6)) and BIC (Eq. (D.7)), we get the relation: $q_{xx} = q_{yx} = q_{xy} = q_{yy} = q$. However, taking the corresponding order of the TE method into account, $k = q_{xx}$, $l = q_{yx}$ for $TE_{Y \rightarrow X}$ and $k = q_{yy}$, $l = q_{xy}$ for $TE_{X \rightarrow Y}$. If we use standard AIC and BIC to estimate the four model orders and set them to the same value, values of parameters k and l will be erroneous. In fact, for most of VAR models, these four model orders are different, and standard AIC and BIC give the maximal value among them. Moreover, if we choose the maximal value for the parameters k and l , TE computational time^v will increase. Focusing on this point, we propose two generalized definitions, namely gAIC and gBIC respectively, as follows:

$$gAIC(\tilde{q}) = N \ln(\det(\hat{\Sigma})) + 2(q_{xx} + q_{yx} + q_{xy} + q_{yy}), \quad (4.85)$$

$$gBIC(\tilde{q}) = N \ln(\det(\hat{\Sigma})) + (q_{xx} + q_{yx} + q_{xy} + q_{yy}) \ln(N) \quad (4.86)$$

where \tilde{q} is a set of estimated orders $\{\hat{q}_{xx}, \hat{q}_{yx}, \hat{q}_{xy}, \hat{q}_{yy}\}$.

^v Computational time can be regarded as the number of times the procedure from Eq. (C.3) to Eq. (C.5) has to be repeated.

It is more reasonable to estimate the four different orders for TE by using the two above definitions. Using gAIC and gBIC to get the four different orders for TE, the computation of TE will be reduced, although the computation time to evaluate the four different orders is increased compared to time needed to compute the order using standard AIC and BIC. For example, using standard AIC and BIC, with $q \in [1, 20]$, procedures from Eq. (D.3) to Eq. (D.5) merely need to be implemented 20 times. Nevertheless, for gAIC and gBIC, this step needs to be implemented 21^4 times since $\{q_{xx}, q_{yx}, q_{xy}, q_{yy}\} \subset [0, 20]^4$, and so, it becomes prohibitive. To solve this problem, we propose a first "greedy" strategy [Shu 2002] to optimize the computation. This technique using gAIC is detailed below (let us note that this technique is the same using gBIC):

- 1) Use standard AIC to obtain a model order, namely the maximum order q_{\max} and select this order to decide the range for gAIC, i.e. $q_{ij} \in [0, q_{\max}]$, $i, j \in \{x, y\}$;
- 2) Consider the values of q_{ij} for the first signal (e.g. q_{xx} and q_{yx}), while setting the values of q_{ij} for the second signal equal to q_{\max} (e.g. q_{xy} and q_{yy}); then implement the procedures from Eq. (D.3) to Eq. (D.5) to get the optimal orders \hat{q}_{xx} and \hat{q}_{yx} by using the definition of Eq. (4.85);
- 3) Consider the values of q_{ij} for the second signal (e.g. q_{xy} and q_{yy}), while setting the values of q_{ij} for the first signal equal to the optimal order \hat{q}_{xx} and \hat{q}_{yx} ; then, implement the procedures from Eq. (D.3) to Eq. (D.5) to get the optimal orders \hat{q}_{xy} and \hat{q}_{yy} by using the definition of Eq. (4.85).

According to the above "greedy" strategy, the computation time is drastically reduced. As an example, the global procedure only needs to be implemented $20 + 2 \times 16^2$ times (20 times for step 1 and 16^2 times for both steps 2 and 3) ($0.27\% ((20 + 2 \times 16^2)/21^4)$) of the computation of gAIC when $q_{\max} = 15$.

4.3.1.2. Selection of radius r

For the selection of the radius r , we take the similar notion as in [Grassberger 1983, Sabesan 2007, Sabesan 2009b]. Firstly, the data are normalized to zero mean and unit variance in order to establish a common radius r in the state space of X and Y . Then, the joint probability in log scale ($\ln C(r)$ obtained as an average of n over $C_{n,r}$ in Eq. (4.79)) is calculated and plotted versus the corresponding radius r in log scale. Finally, the optimal value r^* is chosen in the linear region of the curve $\ln C(r)$ vs $\ln r$ so as to estimate relevant values of TE for this specific value r^* .

4.3.2. Conditional transfer entropy

Obviously, TE is a pairwise analysis based measure. In order to detect and differentiate the direct and indirect relations in multivariate systems, a natural modification is to consider transfer entropy conditional on the knowledge about the environment [Vakorin 2009] and so we focus on a second measure, namely the Conditional Transfer Entropy (CTE). In the following, CTE is developed in the case of three random processes but can be easily extended to multivariate systems.

By considering three random processes X , Y and Z , the relation in Eq. (4.76) can be extended to formalize the absence of information flow from Y to X conditional to Z (for given k , l and s values):

$$P_{X_{n+1}/X_n^{(k)}, Z_n^{(s)}=z_n^{(s)}}(\cdot) = P_{X_{n+1}/X_n^{(k)}, Y_n^{(l)}=y_n^{(l)}, Z_n^{(s)}=z_n^{(s)}}(\cdot). \quad (4.87)$$

The deviation from this assumption can be quantified using the Kullback pseudo-metric, which leads to define the conditional transfer entropy from Y to X conditional to Z :

$$\begin{aligned} CTE_{Y \rightarrow X|Z} = & \int_{\mathbb{R}^{k+l+m+1}} P_{X_{n+1}, X_n^{(k)}, Y_n^{(l)}, Z_n^{(s)}} \left(dx_{n+1}, dx_n^{(k)}, dy_n^{(l)}, dz_n^{(s)} \right) \\ & \times \log_2 \left(\left[\frac{dP_{X_{n+1}/X_n^{(k)}, Y_n^{(l)}, Z_n^{(s)}=z_n^{(s)}}}{dP_{X_{n+1}/X_n^{(k)}, Z_n^{(s)}=z_n^{(s)}}} \right] (x_{n+1}) \right) \end{aligned} \quad (4.88)$$

where the ratio in Eq. (4.88) corresponds to a derivative of the numerator conditional measure with respect to the denominator one. As TE, this measure is not symmetric. It can be estimated from observations (x_n, y_n, z_n) , $n = 1, 2, \dots, N$, using a kernel discrete estimation of $(X_{n+1}, X_n^{(k)}, Y_n^{(l)}, Z_n^{(s)})$ distribution:

$$\hat{P}(x_{n+1}, x_n^{(k)}, y_n^{(l)}, z_n^{(s)}) \propto \sum_m \Theta \left(\left\| \begin{pmatrix} x_{n+1} - x_{m+1} \\ x_n^{(k)} - x_m^{(k)} \\ y_n^{(l)} - y_m^{(l)} \\ z_n^{(s)} - z_m^{(s)} \end{pmatrix} \right\| - r \right) \triangleq C_{n,r}, \quad n = 1, 2, \dots, N-1 \quad (4.89)$$

which depends on a neighborhood size (radius r). Then, it can be used to compute the estimation

$$\widehat{CTE}_{Y \rightarrow X|Z} = \sum_n \hat{P}(x_{n+1}, x_n^{(k)}, y_n^{(l)}, z_n^{(s)}) \log_2 \frac{\hat{P}(x_{n+1} | x_n^{(k)}, y_n^{(l)}, z_n^{(s)})}{\hat{P}(x_{n+1} | x_n^{(k)}, z_n^{(s)})} \quad (4.90)$$

where conditional probabilities in Eq. (4.90) are obtained from the joint probabilities estimated by Eq. (4.89). $\Theta(\cdot)$ is defined by $\Theta(x > 0) = 1$, $\Theta(x \leq 0) = 0$; $\|\cdot\|$ is the sup-norm and summation is performed on $[1, 2, \dots, N-1]$.

Because of Bayes' rule $P(A|B,C) = P(A,B,C)/P(B,C)$, the conditional transfer entropy estimation in Eq. (4.90) can be rewritten as

$$\widehat{CTE}_{Y \rightarrow X|Z} = \sum_n \hat{P}(x_{n+1}, x_n^{(k)}, y_n^{(l)}, z_n^{(s)}) \log_2 \frac{\hat{P}(x_{n+1}, x_n^{(k)}, y_n^{(l)}, z_n^{(s)}) \hat{P}(x_n^{(k)}, z_n^{(s)})}{\hat{P}(x_n^{(k)}, y_n^{(l)}, z_n^{(s)}) \hat{P}(x_{n+1}, x_n^{(k)}, z_n^{(s)})} \quad (4.91)$$

4.3.2.1. Selection of k , l and s

The parameters k , l and s are the orders of the Markov processes for the three coupled processes X , Y and Z and the estimation of the conditional transfer entropy will depend on their values. Similarly as in section 4.3.1.1, we extend the two criteria gAIC and gBIC for the tri-variable VAR processes.

Rewriting Eq. (D.1) with a 3-dimensional VAR process $[z_n(1), z_n(2), z_n(3)]^T = [x_n, y_n, z_n]^T$, we have

$$\begin{cases} x_n = \sum_{i=1}^{q_{xx}} \alpha_{xx}(i) x_{n-i} + \sum_{i=1}^{q_{yx}} \alpha_{yx}(i) y_{n-i} + \sum_{i=1}^{q_{zx}} \alpha_{zx}(i) z_{n-i} + w_{n,x} \\ y_n = \sum_{i=1}^{q_{xy}} \alpha_{xy}(i) x_{n-i} + \sum_{i=1}^{q_{yy}} \alpha_{yy}(i) y_{n-i} + \sum_{i=1}^{q_{zy}} \alpha_{zy}(i) z_{n-i} + w_{n,y} \\ z_n = \sum_{i=1}^{q_{xz}} \alpha_{xz}(i) x_{n-i} + \sum_{i=1}^{q_{yz}} \alpha_{yz}(i) y_{n-i} + \sum_{i=1}^{q_{zz}} \alpha_{zz}(i) z_{n-i} + w_{n,z} \end{cases} \quad (4.92)$$

Obviously, the orders for the index $CTE_{Y \rightarrow X|Z}$ are: $k = q_{xx}$, $l = q_{yx}$ and $s = q_{zx}$. The new definitions of gAIC and gBIC are respectively given by

$$gAIC(\tilde{q}) = N \ln(\det(\hat{\Sigma})) + 2 \sum_{i,j \in \{x,y,z\}} q_{ij} \quad (4.93)$$

$$gBIC(\tilde{q}) = N \ln(\det(\hat{\Sigma})) + \ln(N) \sum_{i,j \in \{x,y,z\}} q_{ij} \quad (4.94)$$

where \tilde{q} is a set with 9 elements $\{\hat{q}_{ij}, i, j \in \{x, y, z\}\}$.

As previously, this approach raises the problem of computation time. Since the number of signals is now equal to 3, the computation time remains prohibitive (for comparison, following the same example as in section 4.3.1.1, the procedure needs to be repeated $20 + 3 \times 16^3$ times when $q_{\max} = 15$) even though we apply the "greedy" strategy as in section 4.3.1.1. Consequently, we propose a second new "greedy" strategy to get an acceptable level of computation.

^{vi} Of course, variables X , Y and Z can be alternately changed to obtain the six available conditional transfer entropies given 3 signals.

The new algorithm using gAIC is detailed below (let us note that the algorithm for gBIC would be similar):

For 3 variables X , Y and Z , we have to determine 9 orders $(q_{ij}, i, j \in \{x, y, z\})$ in Eq. (4.92)) we denote here with only one index $q_i, i \in \{1, 2, \dots, 9\}$. Let us denote a procedure $q_i^{opt} = f_{q_i}(gAIC(\tilde{q}))$, whose output q_i^{opt} is obtained by optimizing the order q_i using the retained criterion (in this case, gAIC):

Initialization

$$V = \{1, 2, \dots, 9\}, \quad q_{\max} = f_q(AIC(q));$$

$$Q = [q_1, q_2, \dots, q_9] = [q_{\max}, q_{\max}, \dots, q_{\max}];$$

Do

choose random i in V ;

$$Q(i) = f_{q_i}(gAIC(Q));$$

$$V = V - \{i\};$$

Until $V = \emptyset$.

According to the above new "greedy" strategy, the computation time is strongly reduced. As an example, when $q_{\max} = 15$, this second new procedure to get gAIC needs to be repeated $20 + 9 \times 16$ times, corresponding to 1.33% of the computation time of the first "greedy" strategy (this first "greedy" strategy proceeded $20 + 3 \times 16^3$ times, which was still too huge) which becomes acceptable.

4.3.2.2. Selection of radius r

For the selection of the radius r , we follow the same methodology as in section 4.3.1.2. To get a common radius r in the state space of X , Y and Z , data are normalized (zero mean and unit variance). Then, the joint probability $\ln C(r)$ is obtained by averaging $C_{n,r}$ over n in Eq. (4.89). If we represent this quantity versus the corresponding radius r in log scale, the optimal value r^* of the radius is obtained in the linear region of the curve $\ln C(r)$ vs $\ln r$. This value r^* guarantees reliable values of TE.

4.4. Discussion

In this chapter, we first introduced a family of well-known methods in the field of Wiener-Granger causality. Secondly, derived from a phase slope based technique, other measures involving different coherence functions are proposed to solve multiple types of relations (direct vs indirect and unidirectional vs bidirectional). Finally, focusing on a more recently developed approach based on information theory, namely the transfer entropy, we proposed a new strategy to select models' orders

so as to improve the estimation and the robustness of transfer entropy. Moreover, to cope with multivariate signals and solve the pairwise analysis issue, the conditional transfer entropy is investigated to detect and differentiate direct and indirect relations in multivariate systems. With a view to further practical implementation (see models and signals in chapter 3), the methodology of this technique and the corresponding strategy to choose the enclosed parameters have been only presented in the trivariate case but can be easily extended to multivariate analysis. All aforementioned techniques are tested in the next chapter on simulated and real signals.

Chapter 5

Experimental results

In this chapter, first of all, we performed Monte Carlo experiments using linear autoregressive (AR) models to test all techniques described in chapter 4 to detect and identify patterns of effective connectivity (e.g. direct or indirect relations, unidirectional or bidirectional influences) leading to a first selection of techniques (section 5.1). Right now, let us indicate that, for each given model, in order to appreciate visually the performance of the different methods, we establish propagation graphs which give a graphic representation of the numerical results. They are obtained following two steps. First of all, we compare the mean value of the index with the corresponding standard deviation. When the ratio is greater than 10, the mean value is considered as meaningful. In a second step, we compare the six mean values of the index under study to decide whether there is a relation or not given a subjective threshold (the "statistical analysis" being performed later). In section 5.2, the selected techniques are tested on nonlinear AR models in the same manner as previously. In a third step (section 5.3), we propose to validate our results on simulated AR signals as well as on physiology-based models using the methodology of evaluation described in chapter 3. Finally, real signals are investigated in section 5.4 before giving some conclusions. The different notations used for each method presented in this chapter are summarized in Appendix E.

5.1. First selection of techniques from linear AR models

Let us recall the expression of the generic linear autoregressive model we considered:

$$\begin{cases} x_1(t) = 0.95\sqrt{2}x_1(t-1) - 0.9025x_1(t-2) + w_1(t) \\ x_2(t) = -0.5x_1(t-1) + 0.25\sqrt{2}x_2(t-1) - \beta x_3(t-3) + w_2(t) \\ x_3(t) = -\alpha x_1(t-2) - 0.5x_2(t-2) - 0.25\sqrt{2}x_3(t-2) + w_3(t) \end{cases} \quad (5.1)$$

where $w_j(t)$, $j = 1, 2, 3$, are realizations of independent white noises W_j with zero mean and unit variance. In order to investigate the effective connectivity, four models including different patterns of causal interactions are extracted from Eq. (5.1) and represented in Fig. 5.1 according to different

combinations of the parameters α and β values. The introduction of the parameter α allows to consider two patterns of causal interactions, either direct relations ($\alpha = 0.5$) or indirect relations ($\alpha = 0$). The parameter β ($\beta = 0.5$) is introduced to model a bidirectional flow between signals 2 and 3. Therefore, the pair of models 1 and 2 shown in Fig. 5.1 is used to detect and distinguish direct or indirect relations from signal 1 to signal 3. The other pair of models, models 3 and 4 shown in Fig. 5.1, is also employed to differentiate and examine direct or indirect relations from signal 1 to signal 3, but, in these two cases, a bidirectional relation is imposed between signals 2 and 3. Moreover, the pair of models 1 and 3, as well as the pair of models 2 and 4, can be considered to inspect the unidirectional and bidirectional relations.

In this section, simulations are carried out 200 times on 2048-point signals and tested by all techniques on all models.

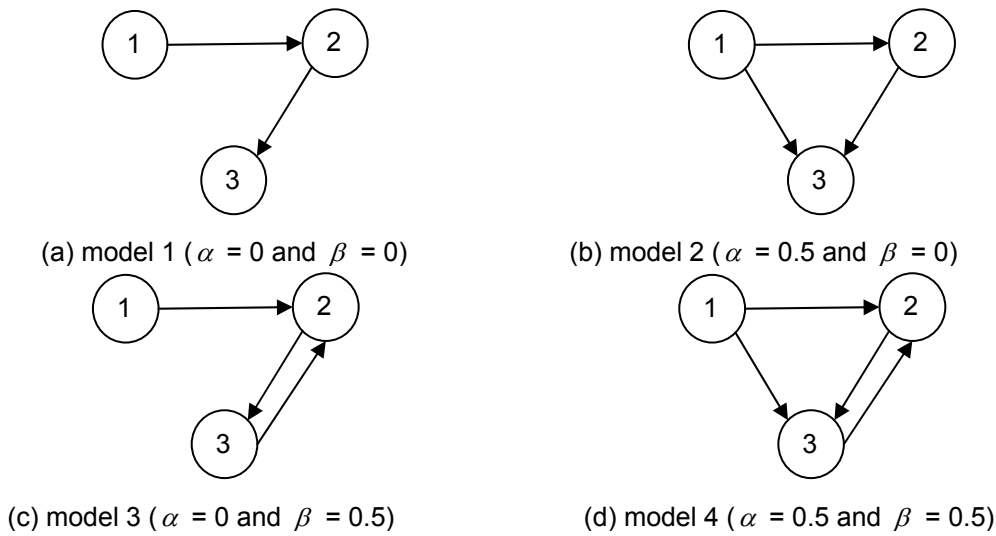


Fig. 5.1 – Schemas for simulated models to generate three signals.

(a): model 1, indirect relation from signal 1 to signal 3 (completely mediated by signal 2) and unidirectional relation from signal 2 to signal 3.

(b): model 2, direct relation from signal 1 to signal 3 and unidirectional relation from signal 2 to signal 3.

(c): model 3, indirect relation from signal 1 to signal 3 (completely mediated by signal 2) and bidirectional relation between signal 2 and signal 3.

(d): model 4, direct relation from signal 1 to signal 3 and bidirectional relation between signal 2 and signal 3.

5.1.1. Granger Causality

In this section, the performance of well-established techniques in the field of Wiener-Granger causality is tested to detect effective connectivity. First of all, WGCI-P and WGCI-C methods are tested in the time domain. Means and standard deviations are derived for each index and presented in a synthesized table for the four linear models (see section 5.1.1.1). Then, according to the previous results, propagation graphs are given for each index (for each index, we have one schema per model). Next, FGCI-P and FGCI-C methods are investigated (see section 5.1.1.2) in the frequency domain. Tables are replaced by figures to follow the frequential behavior of the indices. As previously, schemas of propagation are presented for both techniques and a brief discussion on the results is reported in section 5.1.1.3.

5.1.1.1. *WGCI-P vs WGCI-C*

The results on WGCI-P (given in Eqs. (4.5) and (4.6)) for the four linear models are presented in Tab. 5.1. The first line indicates the mean, and the second line in brackets is the standard deviation (sd). To appreciate visually the performance of the WGCI-P method, we draw on Fig. 5.2 propagation schemas (one per model) which give a graphic representation of the numerical results displayed in Tab. 5.1.

Firstly, as discussed before, WGCI-P is capable of detecting and distinguishing unidirectional and bidirectional relations due to its asymmetry property and Tab. 5.1 confirms this finding. For instance, let us examine the relations between signals 2 and 3 in the pair of models 1 and 3 (the same remark holds when comparing models 2 and 4). For model 1 (or model 2), $WGCI_{2 \rightarrow 3}$ -P leads to an important value whereas $WGCI_{3 \rightarrow 2}$ -P is close to zero indicating that a unidirectional relation exists from 2 to 3. Both $WGCI_{2 \rightarrow 3}$ -P and $WGCI_{3 \rightarrow 2}$ -P give important values for model 3 (and model 4) making it clear that a bidirectional relation exists between signals 2 and 3.

Secondly, to deal with the issue of direct or indirect relations from signal 1 to signal 3, let us compare model 1 and model 2 (or equivalently model 3 and model 4). Obviously, the WGCI-P index cannot distinguish them since $WGCI_{1 \rightarrow 3}$ -P = 0.3783 for model 1 (see Tab. 5.1a) and $WGCI_{1 \rightarrow 3}$ -P = 0.5237 for model 2 (see Tab. 5.1b). In the same manner, for the second pair of models (models 3 and 4), we obtain a similar result: $WGCI_{1 \rightarrow 3}$ -P = 0.3599 for model 3 (see Tab. 5.1c) and $WGCI_{1 \rightarrow 3}$ -P = 0.5014 for model 4 (see Tab. 5.1d).

$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$		$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	0.0014 (0.0008)	0.0013 (0.0008)		$j = 1$	-	0.0014 (0.0007)	0.0015 (0.0011)
$j = 2$	0.7319 (0.0730)	-	0.0037 (0.0031)		$j = 2$	0.7293 (0.0776)	-	0.0457 (0.0126)
$j = 3$	0.3783 (0.0342)	0.5728 (0.0268)	-		$j = 3$	0.5237 (0.0439)	0.3645 (0.0257)	-
(a) model 1					(b) model 2			
$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$		$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	0.0019 (0.0017)	0.0017 (0.0015)		$j = 1$	-	0.0017 (0.0016)	0.0018 (0.0015)
$j = 2$	0.5885 (0.0283)	-	0.2119 (0.0213)		$j = 2$	0.6689 (0.0235)	-	0.3420 (0.0265)
$j = 3$	0.3599 (0.0246)	0.5954 (0.0306)	-		$j = 3$	0.5014 (0.0231)	0.4013 (0.0273)	-
(c) model 3					(d) model 4			

Tab. 5.1 – Results on WGCI-P for the four linear models derived from Eq. (5.1).
The first line indicates the mean and the second line in brackets is the sd.

From Fig. 5.2, the estimated graphs for models 1 and 2 are identical, and so are those estimated for models 3 and 4. This figure confirms that the WGCI-P index fails in distinguishing direct or indirect relations from signal 1 to signal 3 whatever unidirectional or bidirectional relations exist between these two signals.

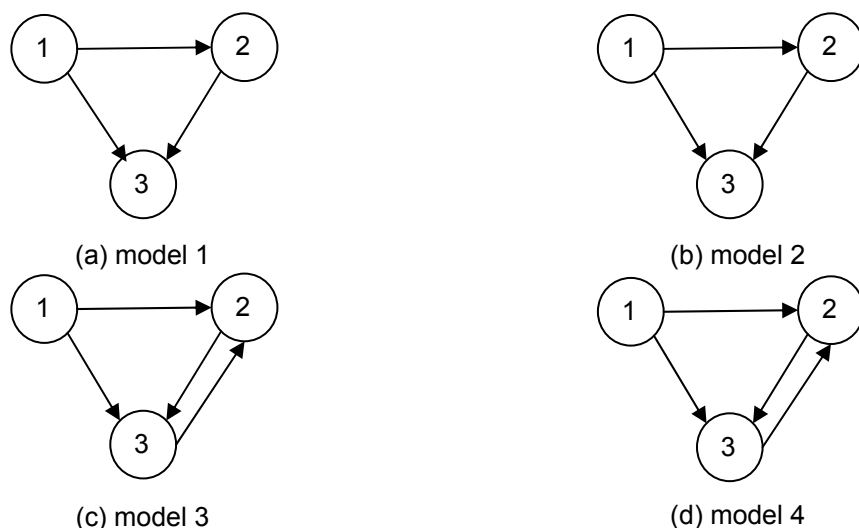


Fig. 5.2 – Estimated graphs from the results on WGCI-P shown in Tab. 5.1 for the four linear models.

Similarly, results on WGCI-C (given in Eq. (4.8)) are summarized in Tab. 5.2 for the four linear models. The corresponding schemas of these results are displayed in Fig. 5.3.

$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$				
$j = 1$	-	0.0012 (0.0005)	0.0010 (0.0003)	$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 2$	0.7154 (0.0309)	-	0.0009 (0.0003)	$j = 1$	-	0.0011 (0.0004)	0.0011 (0.0004)
$j = 3$	0.0012 (0.0004)	0.3344 (0.0230)	-	$j = 2$	0.7283 (0.0299)	-	0.0011 (0.0004)
(a) model 1				$j = 3$	0.4074 (0.0245)	0.3592 (0.0248)	-
				(b) model 2			
$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$	$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	0.0012 (0.0005)	0.0012 (0.0004)	$j = 1$	-	0.0010 (0.0005)	0.0013 (0.0006)
$j = 2$	0.7677 (0.0337)	-	0.3610 (0.0241)	$j = 2$	0.6659 (0.0269)	-	0.3816 (0.0256)
$j = 3$	0.0010 (0.0005)	0.3374 (0.0220)	-	$j = 3$	0.4326 (0.0259)	0.3344 (0.0244)	-
(c) model 3				(d) model 4			

Tab. 5.2 – Results on WGCI-C for the four linear models derived from Eq. (5.1).
The first line indicates the mean and the second line in brackets is the sd.

First of all, from Tab. 5.2, it comes out that the WGCI-C index allows us to identify direct causality relation and distinguish properly the two patterns of causal interactions (direct vs indirect) when comparing model 1 and model 2. As a matter of fact, $WGCI_{1 \rightarrow 3}-C = 0.0012$ (no direct relation) in Tab. 5.2a, whereas $WGCI_{1 \rightarrow 3}-C = 0.4074$ (direct relation) in Tab. 5.2b. The other pair of models (3 and 4) leads to the same conclusion ($WGCI_{1 \rightarrow 3}-C = 0.0010$ in Tab. 5.2c and $WGCI_{1 \rightarrow 3}-C = 0.4326$ in Tab. 5.2d). Consequently, the schemas in Fig. 5.3 reveal the exact propagation graphs for the four linear models and so, we can conclude that, unlike the first index, the WGCI-C index is able to differentiate direct and indirect causal interrelations whatever unidirectional or bidirectional relations exist between the signals.

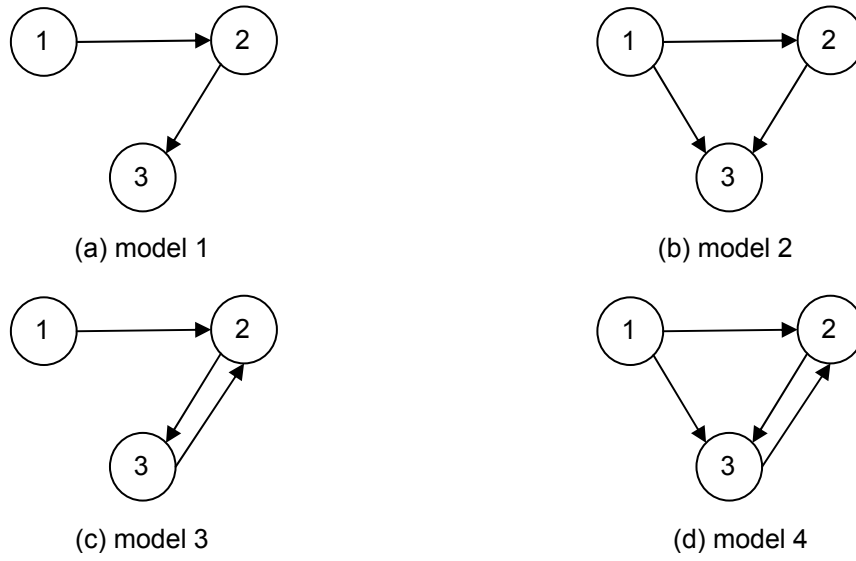


Fig. 5.3 – Estimated graphs from the results on WGCI-C shown in Tab. 5.2 for the four linear models.

5.1.1.2. FGCI-P vs FGCI-C

Results on the mean values of FGCI-P and FGCI-C indices at each frequency f (given in Eqs. (4.18), (4.21) and (4.37)) are represented respectively in Figs. 5.4 and 5.6, and the corresponding propagation graphs in Figs. 5.5 and 5.7.

In Figs. 5.4a and 5.4b, the $FGCI_{1 \rightarrow 3}-P$ index reveals important values around 65 Hz which indicates a causal relation from signal 1 to signal 3, but the nature of this relation is unknown. The $FGCI_{1 \rightarrow 3}-P$ index cannot differentiate the direct or indirect relations from signal 1 to signal 3 comparing models 1 and 2. This conclusion is also true when comparing models 3 and 4 (see Figs. 5.4c and 5.4d). Schemas in Fig. 5.5 reveal this result.

On the other hand, the $FGCI_{1 \rightarrow 3}-C$ index is able to differentiate direct and indirect causal relations from signal 1 to signal 3 whatever unidirectional (e.g. from signal 1 to signal 2) and bidirectional (e.g. between signals 2 and 3) relations exist in the pathway of these two signals according to the results found on the two pairs of models (models 1 and 2 on the one hand, and models 3 and 4 on the other hand) (Figs. 5.6 and 5.7).

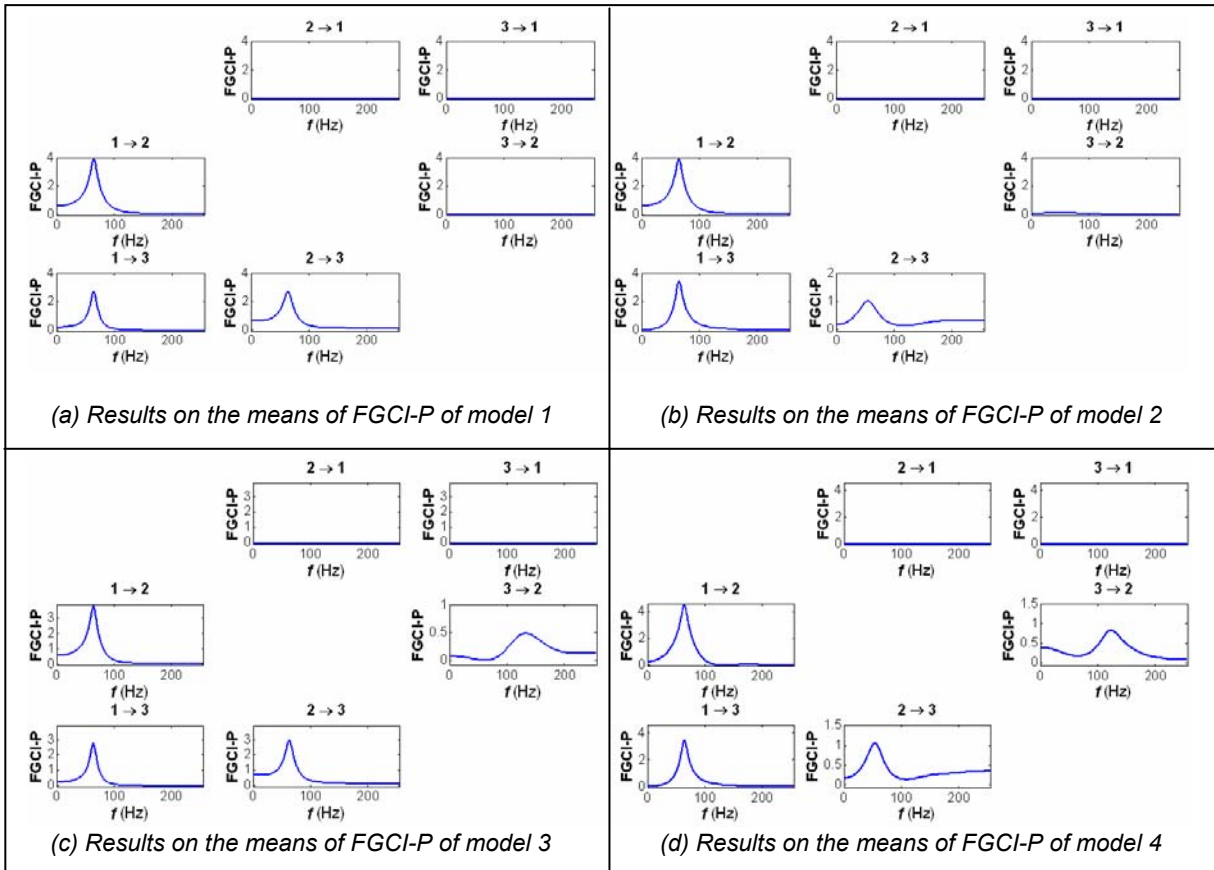


Fig. 5.4 – Mean values of the FGCI-P index at each frequency f .

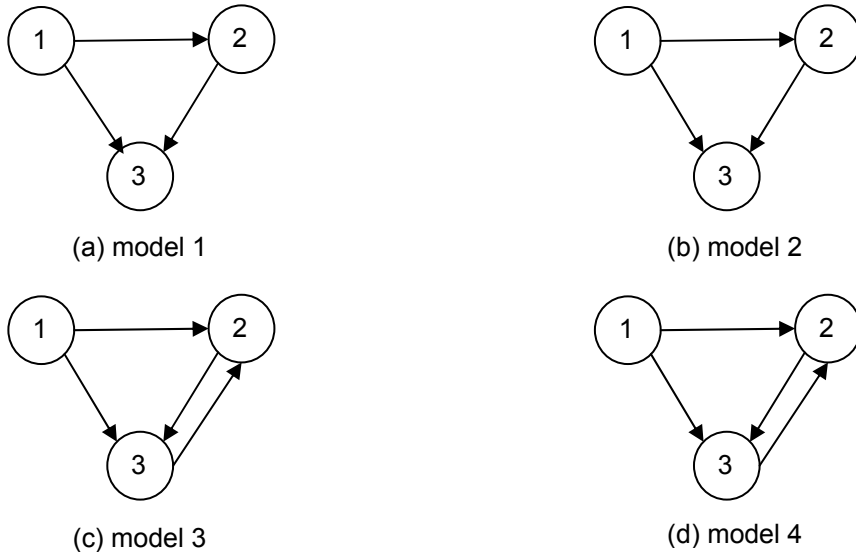


Fig. 5.5 – Estimated graphs from the results on FGCI-P shown in Fig. 5.4 for the four linear models.

5.1.1.3. Discussion and conclusion

From the above analysis, we conclude that (i) in the time domain, the WGCI-C index is able to differentiate direct and indirect causal interrelations whereas the WGCI-P index cannot; (ii) in the frequency domain, the same conclusion holds, *i.e.* the FGCI-C index can distinguish direct and indirect

causal interactions whereas the FGCI-P index fails, (iii) both WGCI-P and WGCI-C (resp. FGCI-P and FGCI-C) indices are capable of identifying and discriminating unidirectional and bidirectional causal interactions in the time (resp. frequency) domain.

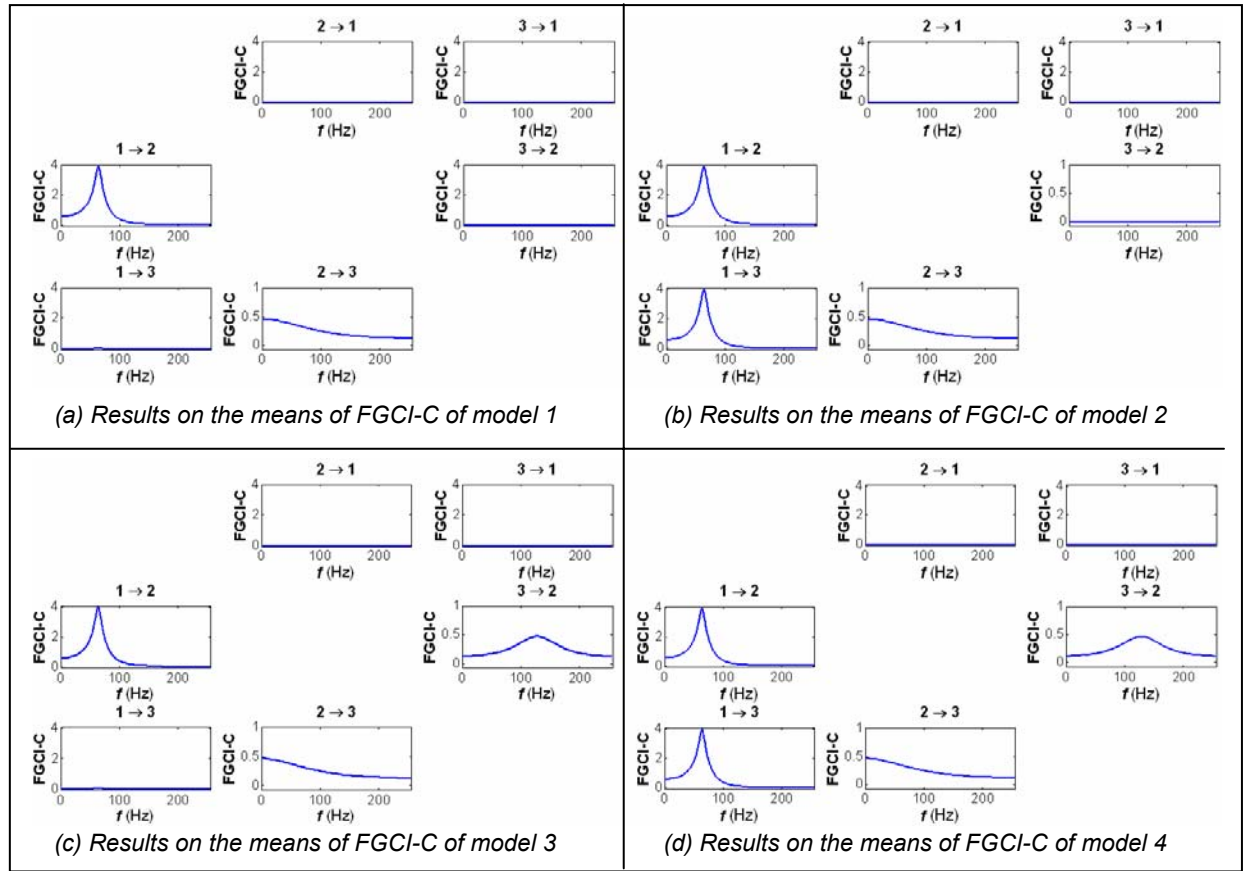


Fig. 5.6 – Mean values of the FGCI-C index at each frequency f .

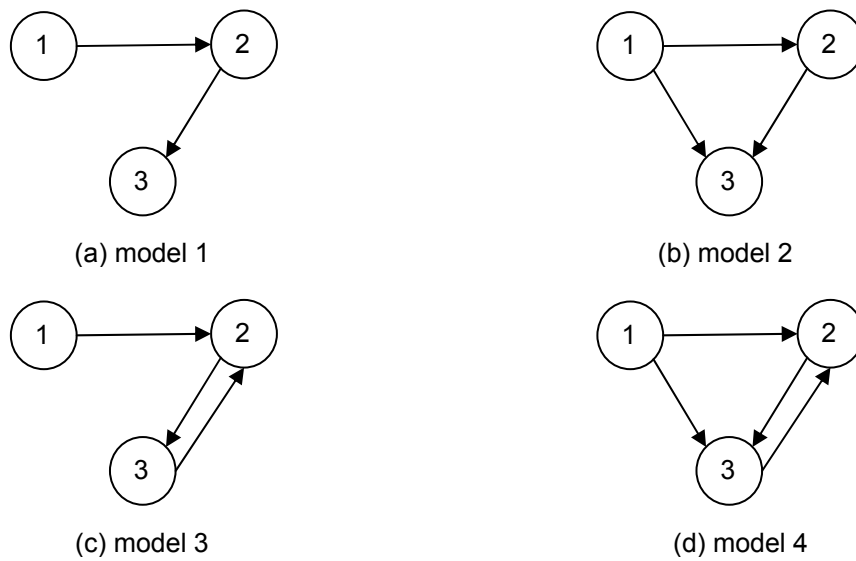


Fig. 5.7 – Estimated graphs from the results on FGCI-C shown in Fig. 5.6 for the four linear models.

5.1.2. Phase slope index and causality index

As we know, the issue of the phase also plays an interesting part in detecting effective connectivity in multivariate system. In this section, we focus on techniques based on the PSI and CI. In section 5.1.2.1, we first compare two measures to estimate signals' spectra involved in the determination of the phase slope index, the first one being based on direct Fourier transforms of signals and the second one using multivariate AR modeling. Then, in section 5.1.2.2, with regard to the issue of direct vs indirect relations in the situation of unidirectional connectivity, we compare the indices given in Eq. (4.52) and (4.62) (the first one involving the OC function and the second one the PC function) through models 1 and 2. Since the OC function between two signals only carries a single direction information (given by the phase itself and its opposite value), it leads to a symmetric index and, consequently, is unable to detect bidirectional flows. To cope with this problem, indices based on DCOH (given in Eq. (4.67)) and DTF (given in Eq. (4.69)) are tested and compared to PSI on models 1 and 3 in section 5.1.2.3. Subsequently, to meet potential direct and indirect relations in bidirectional situations, we recommend to introduce PDC in a new index (given in Eq. (4.71)) compared with DCOH and DTF on models 3 and 4 in section 5.1.2.4, before drawing some conclusions in section 5.1.2.5.

5.1.2.1. PSI-OC

Results on PSI-OC which includes two different spectra-estimated measures on models 1 and 2 are derived and presented in Tab. 5.3. In the following, we denote by PSI-OC(FT) the situation where ordinary coherence is estimated by Fourier transform. In the same way, PSI-OC(AR) denotes the situation where ordinary coherence is estimated by AR modeling. When using FT, spectra are obtained using a sliding window of 64-point length and a 50% overlap. As for AR modeling, it is realized on the whole signal length.

	Model 1 ($\alpha = 0$ and $\beta = 0$)		Model 2 ($\alpha = 0.5$ and $\beta = 0$)	
	PSI-OC(FT)	PSI-OC(AR)	PSI-OC(FT)	PSI-OC(AR)
1 → 2	1.1282 (0.0830)	1.1562 (0.0614)	1.1294 (0.0822)	1.1534 (0.0589)
1 → 3	1.5610 (0.1253)	1.6994 (0.0945)	1.6091 (0.1287)	1.7191 (0.0855)
2 → 3	1.9359 (0.1202)	2.0046 (0.0940)	1.3794 (0.1363)	1.4295 (0.1075)

*Tab. 5.3 – Results on PSI-OC on two different spectra-estimated measures.
The first line indicates the mean and the second line in brackets is the sd.*

First of all, if we compare Tabs. 5.1 and 5.2 with Tab. 5.3, unlike the WGCI-P and WGCI-C indices, the PSI-OC index takes the delay into account: for example, $PSI_{23-OC} > PSI_{12-OC}$ whereas $WGCI_{2 \rightarrow 3} - P < WGCI_{1 \rightarrow 2} - P$ and $WGCI_{2 \rightarrow 3} - C < WGCI_{1 \rightarrow 2} - C$, for models 1 and 2. The last inequalities do not respect the logical order.

Then, we analyze the results in Tab. 5.3 by comparing two different methods (FT versus AR modeling) to estimate the spectra. The results obtained with AR modeling are preferred since (i) the mean values of the index are generally higher with AR modeling, and (ii) the corresponding standard deviations are smaller. Consequently, in the following sections, the estimation of the spectra is obtained by AR modeling.

5.1.2.2. PSI-OC vs CI-PC

Results on PSI-OC and CI-PC for models 1 and 2 are shown in Tab. 5.4, and the corresponding schemas of propagation are given in Fig. 5.8. According to this table, (i) both PSI-OC and CI-PC perfectly point out the flow direction of information among the three signals; (ii) unlike PSI-OC, CI-PC allows distinguishing direct and indirect relations. For example, in the case $1 \rightarrow 3$, when $\alpha = 0$, corresponding to indirect relation, CI_{13} -PC is close to zero, and when $\alpha = 0.5$, corresponding to direct relation, CI_{13} -PC increases significantly, whereas PSI_{13} -OC systematically presents important values whatever α ($\alpha = 0$ or $\alpha = 0.5$). So, the estimated propagation schemas displayed in Fig. 5.8 show clearly that the PSI-OC index does not differentiate the direct or indirect relations from signal 1 to signal 3 whereas the CI-PC index does.

	Model 1 ($\alpha = 0$ and $\beta = 0$)		Model 2 ($\alpha = 0.5$ and $\beta = 0$)	
	PSI-OC	CI-PC	PSI-OC	CI-PC
$1 \rightarrow 2$	1.1562 (0.0614)	0.8170 (0.0533)	1.1534 (0.0589)	0.7377 (0.0422)
$1 \rightarrow 3$	1.6994 (0.0945)	0.0056 (0.0023)	1.7191 (0.0855)	1.1518 (0.0816)
$2 \rightarrow 3$	2.0046 (0.0940)	1.2944 (0.0864)	1.4295 (0.1075)	1.2889 (0.0886)

Tab. 5.4 – Results on PSI-OC and CI-PC for models 1 and 2.
The first line indicates the mean and the second line in brackets is the sd.

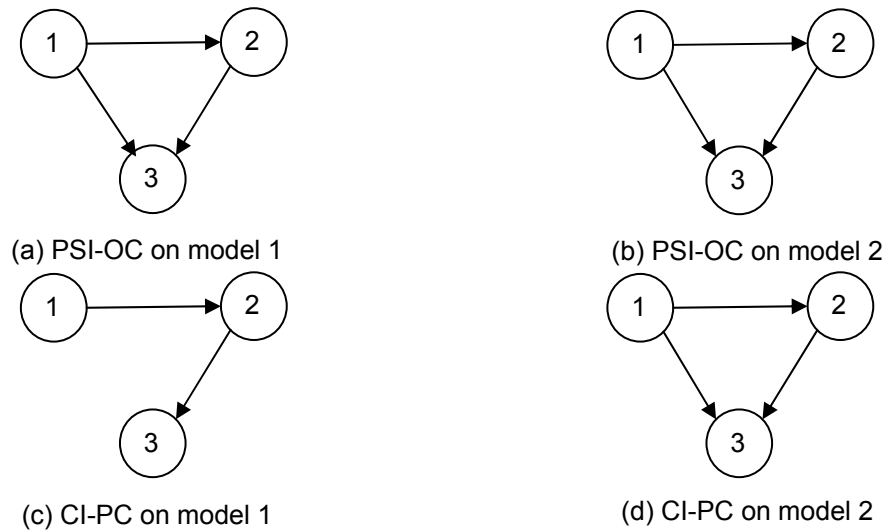


Fig. 5.8 – Estimated graphs from the results on PSI-OC and CI-PC shown in Tab. 5.4.

5.1.2.3. PSI-OC vs CI-DCOH and CI-DTF

Results on PSI-OC, CI-DCOH and CI-DTF for models 1 and 3 are summarized in Tab. 5.5. The corresponding propagation schemas are shown in Fig. 5.9. Examining Tabs. 5.5a and 5.5b, PSI-OC points out the correct information flow when there is only unidirectional causality, e.g. between signals 1 and 2. On the other hand, PSI-OC fails when bidirectional causality exists, e.g. between signals 2 and 3. Distinctly, from Fig. 5.9b, PSI-OC is deemed unable to deal with bidirectional relations, in so far as it considers only a unidirectional causality relation from signal 2 to signal 3.

$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	-1.1562 (0.0614)	-1.6994 (0.0945)
$j = 2$	1.1562 (0.0614)	-	-2.0046 (0.0940)
$j = 3$	1.6994 (0.0945)	2.0046 (0.0940)	-
(a) PSI-OC on model 1			
$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	-0.8535 (0.0822)	-1.4417 (0.1289)
$j = 2$	0.8535 (0.0822)	-	-0.8140 (0.1246)
$j = 3$	1.4417 (0.1289)	0.8140 (0.1246)	-
(b) PSI-OC on model 3			
$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	0.0064 (0.0056)	0.0117 (0.0115)
$j = 2$	2.8975 (0.0628)	-	0.0546 (0.0344)
$j = 3$	3.4820 (0.1290)	3.0424 (0.0889)	-
(c) CI-DCOH on model 1			
$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	0.0331 (0.0259)	0.0301 (0.0203)
$j = 2$	2.6497 (0.0767)	-	1.6491 (0.1418)
$j = 3$	3.2084 (0.1363)	2.8943 (0.0942)	-
(d) CI-DCOH on model 3			
$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	0.0066 (0.0060)	0.0034 (0.0030)
$j = 2$	2.9474 (0.0484)	-	0.0053 (0.0043)
$j = 3$	3.5168 (0.0858)	1.0387 (0.0683)	-
(e) CI-DTF on model 1			
$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	0.0096 (0.0066)	0.0122 (0.0081)
$j = 2$	2.6840 (0.0578)	-	1.6273 (0.0905)
$j = 3$	3.2700 (0.1109)	1.1088 (0.0636)	-
(f) CI-DTF on model 3			

Tab. 5.5 – Results on PSI-OC, CI-DCOH and CI-DTF for models 1 and 3.

The first line indicates the mean and the second line in brackets is the sd.

If we analyze results on CI-DCOH, we find that, in the case of the first model, CI_{23} -DCOH = 3.0424 and CI_{32} -DCOH = 0.0546 (Tab. 5.5c) and therefore reach to the conclusion that a unidirectional relation is detected from signal 2 to signal 3 (see also Fig. 5.9c). Concerning the third model, we find that CI_{23} -DCOH = 2.8943 and CI_{32} -DCOH = 1.6491 (Tab. 5.5d), which reveals the

effective bidirectional relation between signals 2 and 3 (see also Fig. 5.9d). The same conclusion holds when analyzing CI-DTF, which also appears as a relevant index to detect unidirectional and bidirectional relations (Figs. 5.9e and 5.9f).

Nevertheless, both CI-DCOH and CI-DTF give some important values for the causality interaction from signal 1 to signal 3 whatever model 1 or 3. Owing to the fact that only indirect relation exists from signal 1 to signal 3, it seems that CI-DCOH and CI-DTF are unable to make a distinction between direct and indirect relations. This problem is discussed in the next section and solved thanks to CI-PDC.

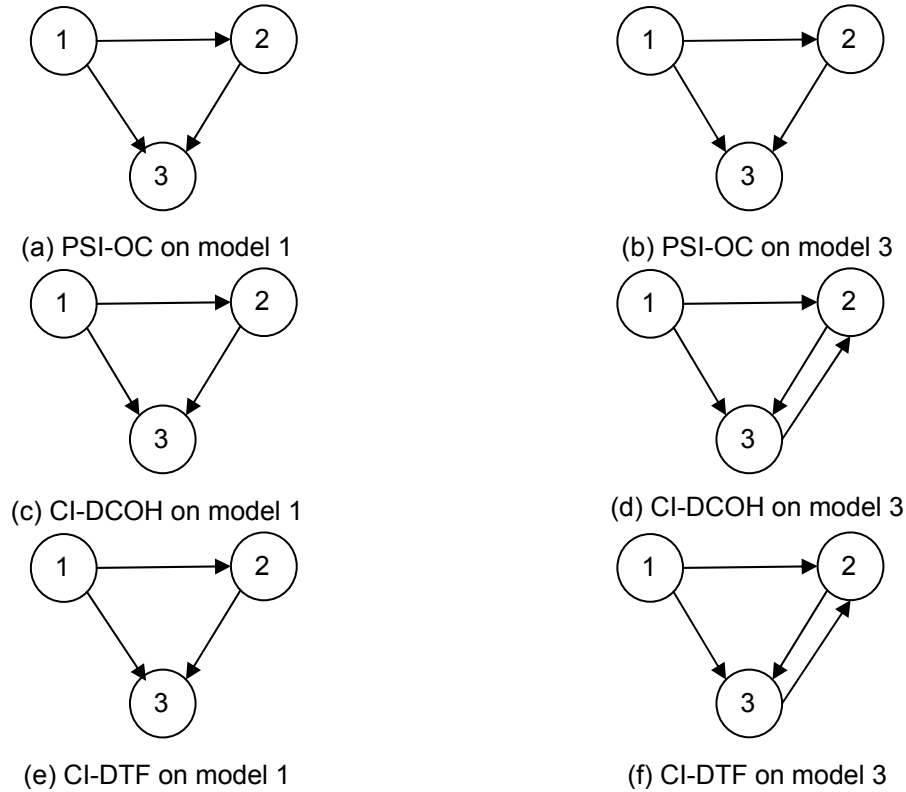


Fig. 5.9 – Estimated graphs from the results on PSI-OC, CI-DCOH and CI-DTF shown in Tab. 5.5.

5.1.2.4. CI-DCOH and CI-DTF vs CI-PDC

Results on CI-DCOH, CI-DTF and CI-PDC for models 3 and 4 are presented in Tab. 5.6 and the corresponding schemas are shown in Fig. 5.10.

The CI-DCOH index allows for pointing out unidirectional and bidirectional relations (see Tabs. 5.6a and 5.6b either for model 3 or for model 4). However, when two different modes of interaction exist between 1 and 3, *i.e.* only indirect relations in the case of model 3, and direct and indirect relations in the case of model 4 (signal 1 connects to signal 3 via two distinct pathways), CI_{13} -DCOH reveals comparable values and, consequently, it becomes impossible to discriminate between direct and indirect pathway. Likewise, the CI-DTF index is successful in indicating unidirectional and bidirectional relations but fails in distinguishing direct and indirect relations from signal 1 to signal 3 for models 3 and 4 (see Tabs. 5.6c and 5.6d).

$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$		$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	0.0331 (0.0259)	0.0301 (0.0203)		$j = 1$	-	0.0323 (0.0227)	0.0279 (0.0192)
$j = 2$	2.6497 (0.0767)	-	1.6491 (0.1418)		$j = 2$	3.3775 (0.1230)	-	1.9630 (0.1161)
$j = 3$	3.2084 (0.1363)	2.8943 (0.0942)	-		$j = 3$	3.2975 (0.1177)	1.7631 (0.1247)	-
(a) CI-DCOH on model 3					(b) CI-DCOH on model 4			
$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$		$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	0.0096 (0.0066)	0.0122 (0.0081)		$j = 1$	-	0.0090 (0.0065)	0.0111 (0.0068)
$j = 2$	2.6840 (0.0578)	-	1.6273 (0.0905)		$j = 2$	3.5014 (0.0639)	-	1.6029 (0.0890)
$j = 3$	3.2700 (0.1109)	1.1088 (0.0636)	-		$j = 3$	3.3417 (0.0887)	1.0504 (0.0688)	-
(c) CI-DTF on model 3					(d) CI-DTF on model 4			
$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$		$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	0.0054 (0.0048)	0.0049 (0.0045)		$j = 1$	-	0.0057 (0.0047)	0.0053 (0.0051)
$j = 2$	0.9535 (0.0330)	-	1.9960 (0.1029)		$j = 2$	0.6033 (0.0322)	-	1.9941 (0.1051)
$j = 3$	0.0052 (0.0050)	1.3280 (0.0640)	-		$j = 3$	1.2069 (0.0663)	1.3370 (0.0739)	-
(e) CI-PDC on model 3					(f) CI-PDC on model 4			

*Tab. 5.6 – Results on CI-DCOH, CI-DTF and CI-PDC for models 3 and 4.
The first line indicates the mean and the second line in brackets is the sd.*

Like the CI-DCOH and CI-DTF indices, the CI-PDC index succeeds in detecting unidirectional and bidirectional relations (Tabs. 5.6e and 5.6f). By contrast, when there is only indirect flow from signal 1 to signal 3 (for $\alpha = 0$, in model 3), CI_{13} -PDC remains close to zero (Tab. 5.6e), and, when there is direct relation from 1 to 3 (for $\alpha = 0.5$, in model 4), CI_{13} -PDC increases significantly (Tab. 5.6f). Consequently, the CI-PDC index outperforms CI-DCOH and CI-DTF indices, since these two indices always display a non negligible value when they are used to test the relation from signal 1 to signal 3, whatever the value of α . Therefore, the CI-PDC index resolves the existence of direct and indirect connections between pairs of signals by differentiating these kinds of connection. Furthermore, even if the CI-PDC and CI-DTF indices cannot reveal the delays between signals in the bidirectional case, one can at least infer the value of the delays' ratio between signals 2 and 3 (a 2-time delay from signal 2 to signal 3 and a 3-time delay from signal 3 to signal 2). As a matter of fact, the ratios

$CI_{32}\text{-PDC}/CI_{23}\text{-PDC}$ and $CI_{32}\text{-DTF}/CI_{23}\text{-DTF}$ are always around 3/2 (in model 3 or 4) whereas the ratio $CI_{32}\text{-DCOH}/CI_{23}\text{-DCOH}$ varies much more (around 0.57 for model 3 and 1.11 for model 4). The conclusions drawn before can be easily visualized in Fig. 5.10.

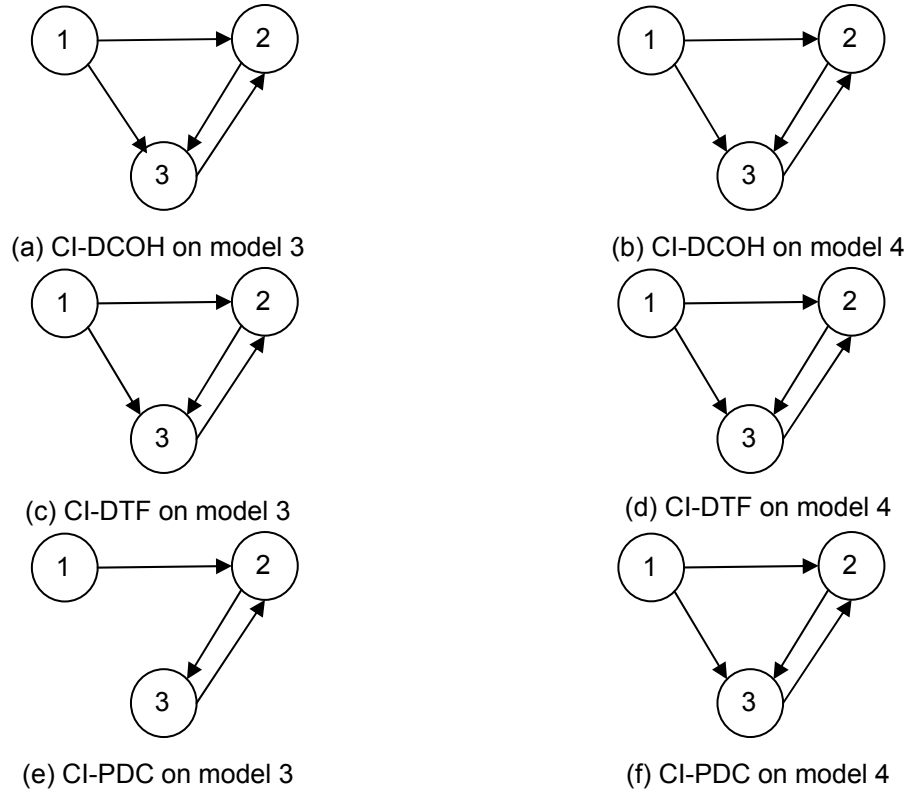


Fig. 5.10 – Estimated graphs from the results on CI-DCOH, CI-DTF and CI-PDC shown in Tab. 5.6.

5.1.2.5. Discussion and conclusion

In this section, we first compared two estimators of signals' spectra (Fourier transform versus AR modeling). The most efficient one – AR modeling – (which is not surprising when applied on simulated AR data) is then kept when estimating all connectivity indices. In the case of unidirectional flow, the phase slope index originally proposed by Nolte is first improved by introducing the partial coherence. In a second step, to cope with the issue of unidirectional and bidirectional effective connectivity, two new indices, one based on the directed coherence and the other on the directed transfer function are investigated. Finally, a new index based on partial directed coherence is proposed to detect and distinguish various types of effective connectivity (direct vs indirect and unidirectional vs bidirectional), and experimental results prove its relevance.

5.1.3. Transfer entropy and conditional transfer entropy

In this section, a family of techniques based on information theory is tested to investigate various patterns of effective connectivity. Transfer entropy which makes use of time series data to quantify the amount of information transferred from one process to another becomes a prominent method to estimate causal interaction in time series analysis and is often proposed to detect effective connectivity in neuroscience. In section 5.1.3.1, we first study the performance of TE (given in Eq.

(4.81)) by means of two improved aspects of parameter selections. Afterwards, TE with improved parameter selections is tested on the four linear models. Finally, in contrast to TE which is a pairwise-analysis method, we consider a trivariate analysis investigating CTE on the four linear models in section 5.1.3.2 before drawing some conclusions in section 5.1.3.3.

5.1.3.1. TE

As introduced in section 4.3.1, internal parameters play a crucial role in TE estimation. Therefore, in the following, experiments are conducted according to two critical parameters: (i) the orders k and l of the two coupled processes under study, (ii) the radius r in the state space for the estimation of multidimensional joint probabilities. In order to explain how to choose and optimize these parameters, we introduce a simple bivariate AR model given by Eq. (5.2) since TE analyzes causal relations by pairs (of course, after an optimization of these parameters, TE will be computed on the linear trivariate AR models given by Eq. (5.1)). The bivariate AR model we choose is as follows:

$$\begin{cases} x_1(t) = 0.95\sqrt{2}x_1(t-1) - 0.9025x_1(t-2) + w_1(t) \\ x_2(t) = -0.5x_1(t-1) + 0.25\sqrt{2}x_2(t-1) + w_2(t) \end{cases} \quad (5.2)$$

where $w_j(t)$, $j = 1, 2$, are realizations of independent white noises W_j with zero mean and unit variance.

1) Selection of the orders of the two coupled processes

To compare the performance of the four selection order criteria (AIC, BIC, gAIC and gBIC), we present in Tab. 5.7 the number of times that the model's orders fit the expected (optimal) orders over the 200 trials.

	$\hat{q}_{11} = q_{11}(2)$	$\hat{q}_{21} = q_{21}(0)$	$\hat{q}_{12} = q_{12}(1)$	$\hat{q}_{22} = q_{22}(1)$
AIC	190/200	0/200	0/200	0/200
BIC	200/200	0/200	0/200	0/200
gAIC	198/200	184/200	180/200	170/200
gBIC	200/200	200/200	198/200	199/200

Tab. 5.7 – Number of times the orders estimated by the four different measures equal the expected orders of the model given by Eq. (5.2).

As discussed in Section 4.3.1.1, the expected order values of the model described by Eq. (5.2) are $q_{11} = 2$, $q_{21} = 0$, $q_{12} = 1$ and $q_{22} = 1$. In this way, considering TE, we must obtain: $k = q_{11} = 2$, $l = q_{21} = 0$ for $TE_{2 \rightarrow 1}$ and $k = q_{12} = 1$, $l = q_{22} = 1$ for $TE_{1 \rightarrow 2}$. As recalled before, AIC and BIC return only one order value which must correspond to the maximum order of the model. In TE approach, the four estimated orders using AIC and BIC are set to this value and, in the present case, theoretically we must obtain the following estimated values $\hat{q}_{11} = \hat{q}_{21} = \hat{q}_{12} = \hat{q}_{22} = 2$ (see Tab. 5.7). In this manner, in the model described by Eq. (5.2) three parameters out of four cannot be properly estimated (since these parameters differ from 2). The gAIC measure is a reasonable estimator since most of the

estimated orders are the optimal ones (for example, 184 correct estimations over a total of 200 (92%) for \hat{q}_{21}). Obviously, in this experiment, the gBIC measure behaves the best (almost 100% of accuracy estimation for the four parameters).

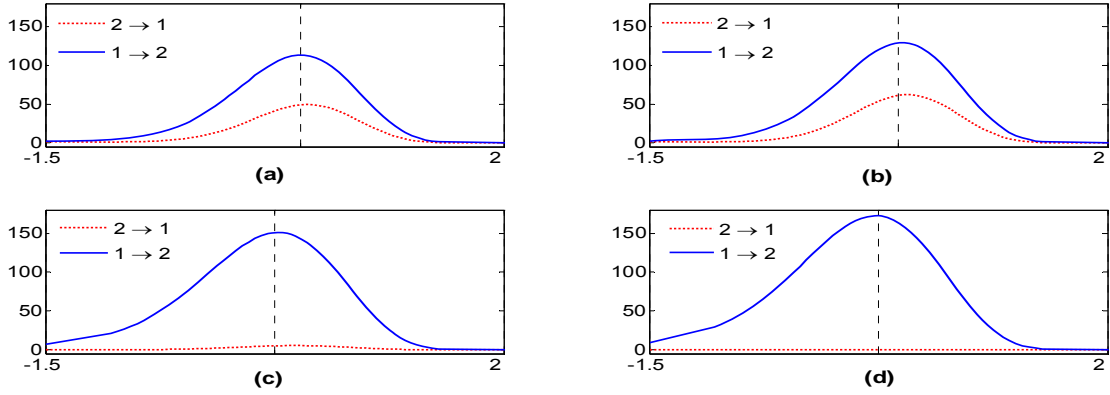


Fig. 5.11 – Mean values of the TE index on the model derived from Eq. (5.2).

The x-axis is the radius r (in \ln scale), and the y-axis is TE (bits).

(a) The orders are estimated by AIC.

(b) The orders are estimated by BIC.

(c) The orders are estimated by gAIC.

(d) The orders are estimated by gBIC.

Given the orders estimated by the four techniques, we compute the averaged values for TE over the 200 trials. Results are shown in Fig. 5.11 for different values of the radius r (from $e^{-1.5}$ to $e^{2.0}$ with a step of $e^{0.05}$). Using gBIC, $TE_{1 \rightarrow 2}$ displays the largest amplitude in some range of $\ln r$ (around the vertical dotted line) when a real flow exists from 1 to 2 while $TE_{2 \rightarrow 1}$ is zero when there is no information flow from 2 to 1 (Fig. 5.11d). Using standard AIC and BIC (see Figs. 5.11a and 5.11b), we note some non-zero values for $TE_{2 \rightarrow 1}$ due to some erroneous model orders: the estimated order $l = \hat{q}_{x_2 x_1}$ is always 2 which is not the right order (it must be 0). Using gAIC, $TE_{1 \rightarrow 2}$ is close to $TE_{1 \rightarrow 2}$ using gBIC, while $TE_{2 \rightarrow 1}$ is mostly close to zero (see Fig. 5.11c) since only a few estimated orders are not the optimal ones.

2) Selection of the radius r

As discussed in section 4.3.1.2, a common value of radius r is desirable for the computation of the multi-dimensional joint probabilities. To establish this common radius r in the state space of X_1 and X_2 , the data are first normalized to zero mean and unit variance. Then, we compute all multi-dimensional joint probabilities needed in Eq. (4.81) to derive TE (using the orders estimated by each technique). In order to select the radius, we compute an averaged value of the joint probability defined by Eq. (4.79) and choose the specific radius r^* in the linear region of the curve $\ln C(r)$ vs $\ln r$ where C_r is the mean value of $C_{n,r}$ (see Fig. 5.12a). In Fig. 5.12b, we plot TE (bits) vs r . At the specific radius r^* , $TE_{1 \rightarrow 2}$ achieves the maximum value when information flow exists from signal 1 to signal 2,

and $TE_{2 \rightarrow 1}$ remains zero when there is no information flow from signal 2 to signal 1. Afterwards, the means and sd of the corresponding TE values with the orders estimated by the four criteria at this radius r^* are summarized in Tab. 5.8.

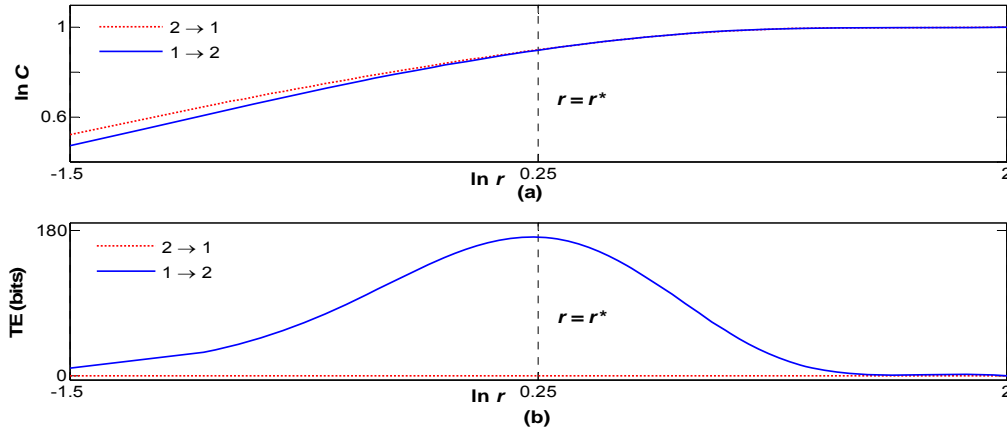


Fig. 5.12 – Transfer entropy with the orders estimated by gBIC between signals 1 and 2 for the model given by Eq. (5.2).

(a) Plot of the mean of $C_{n,r}$ with respect to n vs r (in \ln scale), where C_r and r denote average joint probability and radius in the state space respectively.

(b) Plot of TE (bits) vs radius r (in \ln scale), for $1 \rightarrow 2$ flow direction (solid blue line), and $2 \rightarrow 1$ flow direction (dotted red line).

TE	AIC	BIC	gAIC	gBIC
$1 \rightarrow 2$	113.6144 (28.8490)	129.3734 (15.7382)	150.8441 (39.0125)	172.2545 (23.9073)
$2 \rightarrow 1$	49.1832 (24.0837)	62.2040 (13.5001)	5.2758 (15.4259)	0.0000 (0.0000)

Tab. 5.8 – Estimated TE with the orders estimated by the four criteria for the model of Eq. (5.2). The first line indicates the mean and the second line in brackets is the sd.

For sake of clarity, the notation $TE_{1 \rightarrow 2}(C)$ represents results of $TE_{1 \rightarrow 2}$ corresponding to criterion C. From Tab. 5.8, we observe that $TE_{1 \rightarrow 2}(\text{gBIC})$ gives the greatest values when the direction of information flow is from 1 to 2, and $TE_{2 \rightarrow 1}(\text{gBIC})$ is zero when no information flow exists from 2 to 1 thanks to the perfect estimation of the orders. In contrast, $TE_{2 \rightarrow 1}(\text{AIC})$, $TE_{2 \rightarrow 1}(\text{BIC})$, and $TE_{2 \rightarrow 1}(\text{gAIC})$ still provide some non-zero values due to erroneous estimated-orders values.

From the above analysis, the gBIC technique seems to be the most prominent method to evaluate the model order for the more accurate and robust estimation of TE. In order to confirm this conclusion, results on model 1 (Eq. (5.1)) are presented in Tab. 5.9. From this table, we find that $TE_{i \rightarrow j}(\text{gBIC})$ gives the greatest values when the direction of information flow from signal i to signal j (e.g. $1 \rightarrow 2$, $1 \rightarrow 3$ and $2 \rightarrow 3$) exists, and $TE_{i \rightarrow j}(\text{gBIC})$ is zero when there is no information flow (e.g. $2 \rightarrow 1$, $3 \rightarrow 1$ and $3 \rightarrow 2$).

$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$		$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	49.1832 (24.0837)	4.3685 (2.3604)		$j = 1$	-	62.2040 (13.5001)	7.8292 (1.4785)
$j = 2$	113.6144 (28.8490)	-	5.4181 (1.3790)		$j = 2$	129.3734 (15.7382)	-	10.9475 (3.2243)
$j = 3$	32.5233 (4.7463)	29.1554 (3.7699)	-		$j = 3$	38.2687 (2.8332)	53.8152 (7.7682)	-
(a) $TE_{i \rightarrow j}$ (AIC)					(b) $TE_{i \rightarrow j}$ (BIC)			
$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$		$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	5.2758 (15.4259)	2.5744 (8.0493)		$j = 1$	-	0.0000 (0.0000)	0.0000 (0.0000)
$j = 2$	150.8441 (39.0125)	-	0.6079 (1.3124)		$j = 2$	172.2545 (23.9073)	-	0.0000 (0.0000)
$j = 3$	41.5814 (7.7983)	72.4007 (11.7379)	-		$j = 3$	44.0241 (5.5336)	77.1240 (4.0925)	-
(c) $TE_{i \rightarrow j}$ (gAIC)					(d) $TE_{i \rightarrow j}$ (gBIC)			

Tab. 5.9 – Results on TE with the orders estimated by the four criteria for model 1 (in Eq. (5.1)).

The first line indicates the mean and the second line in brackets is the sd.

(a) The orders are estimated by AIC. (b) The orders are estimated by BIC.

(c) The orders are estimated by gAIC. (d) The orders are estimated by gBIC.

$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$		$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	0.0000 (0.0000)	0.0000 (0.0000)		$j = 1$	-	0.0000 (0.0000)	0.0000 (0.0000)
$j = 2$	172.2545 (23.9073)	-	0.0000 (0.0000)		$j = 2$	173.6717 (23.7270)	-	0.0000 (0.0000)
$j = 3$	44.0241 (5.5336)	77.1240 (4.0925)	-		$j = 3$	67.3186 (6.7619)	72.0084 (3.7393)	-
(a) model 1					(b) model 2			
$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$		$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	0.0000 (0.0000)	0.0000 (0.0000)		$j = 1$	-	0.0000 (0.0000)	0.0000 (0.0000)
$j = 2$	43.8141 (2.6655)	-	11.4119 (1.8512)		$j = 2$	48.8013 (1.8513)	-	21.1229 (1.6838)
$j = 3$	21.4008 (1.6229)	34.3956 (4.3808)	-		$j = 3$	27.7992 (2.4871)	19.8090 (2.0137)	-
(c) model 3					(d) model 4			

Tab. 5.10 – Results on TE with the orders estimated by gBIC on the four linear models (in Eq. (5.1)).

The first line indicates the mean and the second line in brackets is the sd.

In the following, we test TE on the four linear models using gBIC. Results are given in Tab. 5.10 and the corresponding graphs in Fig. 5.13. What we can conclude on these results is that (i) unidirectional and bidirectional relations are correctly established, (ii) direct and indirect relations are ambiguous. This issue can be solved by conditional transfer entropy and the corresponding experimental results are given in the following section.

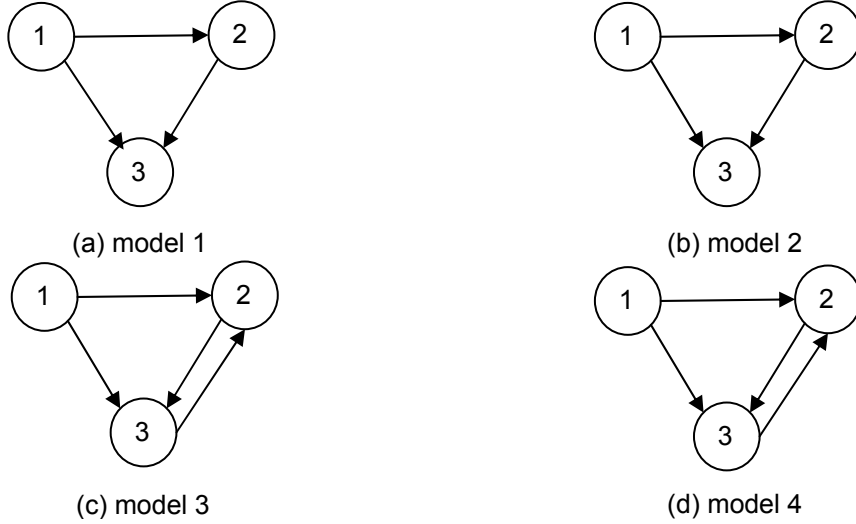


Fig. 5.13 – Estimated graphs from the results on TE shown in Tab. 5.10 for the four linear models.

5.1.3.2. CTE

As discussed in sections 4.3.2.1 and 4.3.2.2, we follow the same methodology to choose the specific r^* in the linear region of the curve $\ln C(r)$ vs $\ln r$. An example is shown in Fig. 5.14. The second new "greedy" strategy is implemented to reduce the computation time in the orders selection for gAIC and gBIC (Eqs. (4.93) and (4.94)). In the following, the performance of the four criteria (AIC, BIC, gAIC and gBIC) are compared on CTE index. The results are exhibited in the following figures and tables (the next 4 pages).

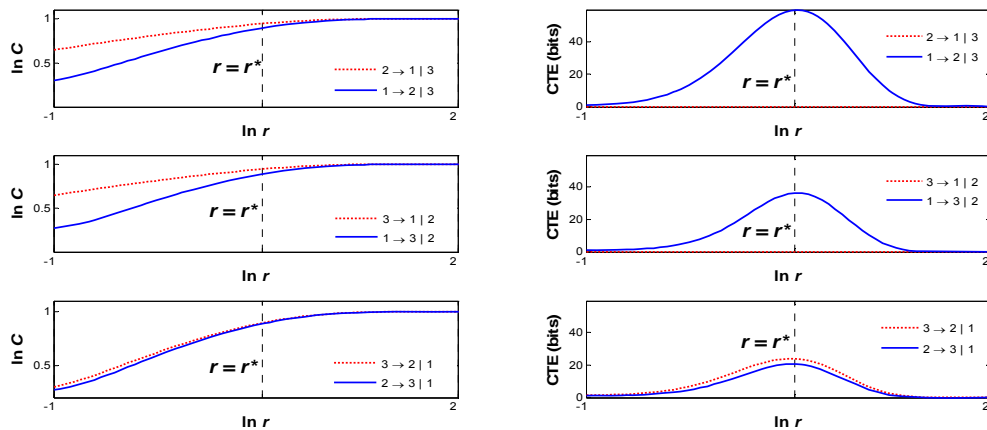


Fig. 5.14 – Results on CTE with the orders estimated by gBIC for model 4 derived from Eq. (5.1).
Left panel: Plot of the mean of $C_{n,r}$ with respect to n vs r (in \ln scale), where C_r and r denote average joint probability and radius in the state space respectively.
Right panel: Plot of CTE (bits) vs radius r (in \ln scale).

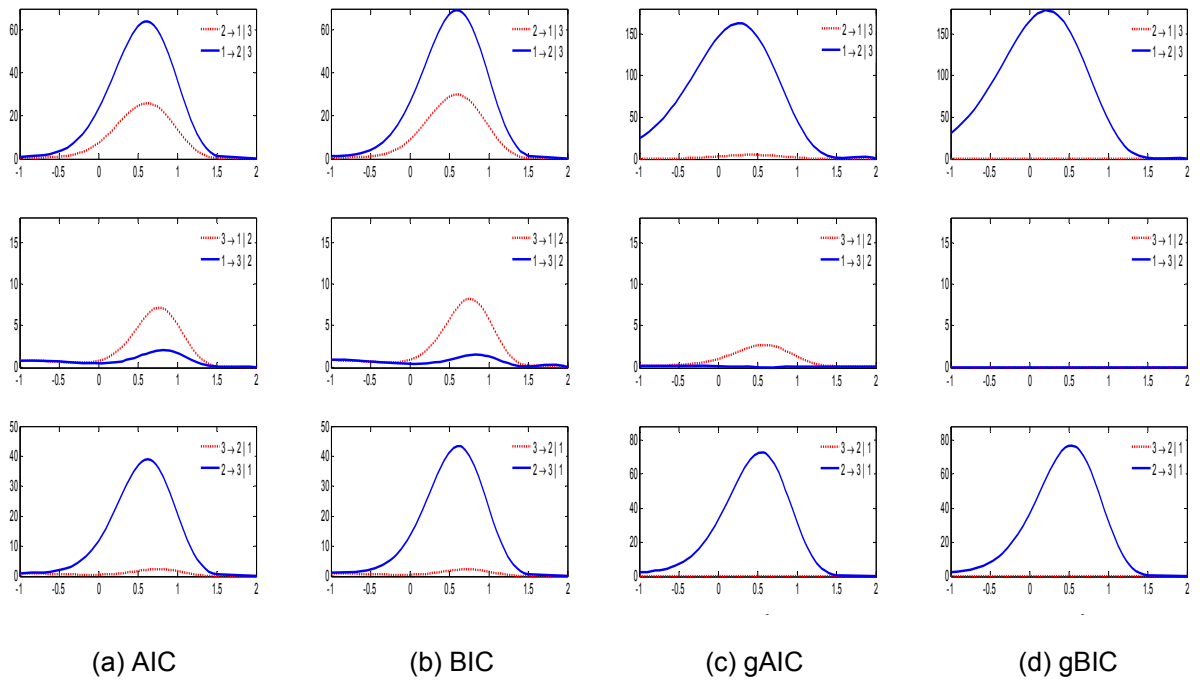


Fig. 5.15 – Results on CTE (bits) with the orders estimated by the four criteria for model 1 (in Eq. (5.1)).
For each criterion, the six available CTE are plotted in one column.

$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$				
$j = 1$	-	26.0646 (8.5208)	6.1041 (2.5261)	$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 2$	64.1777 (11.6090)	-	2.0800 (0.7455)	$j = 1$	-	29.9656 (2.8904)	7.1568 (1.4098)
$j = 3$	1.5866 (1.1851)	38.9012 (9.7438)	-	$j = 2$	69.3401 (3.4089)	-	2.0976 (0.7654)
(a) AIC				$j = 3$	1.1098 (0.6169)	43.3415 (2.7956)	-
				(b) BIC			
$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$	$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	4.5694 (13.7305)	1.7984 (5.7885)	$j = 1$	-	0.0000 (0.0000)	0.0000 (0.0000)
$j = 2$	162.8025 (30.3258)	-	0.0000 (0.0000)	$j = 2$	178.0639 (8.8438)		0.0000 (0.0000)
$j = 3$	0.0135 (0.1855)	56.7869 (10.0513)	-	$j = 3$	0.0011 (0.0188)	56.7329 (3.9738)	-
(c) gAIC				(d) gBIC			

Tab. 5.11 – Results on CTE at specific r^* for model 1 (in Eq. (5.1)).
The first line indicates the mean and the second line in brackets is the sd.

(a) The orders are estimated by AIC.

(b) The orders are estimated by BIC.

(c) The orders are estimated by gAIC.

(d) The orders are estimated by gBIC.

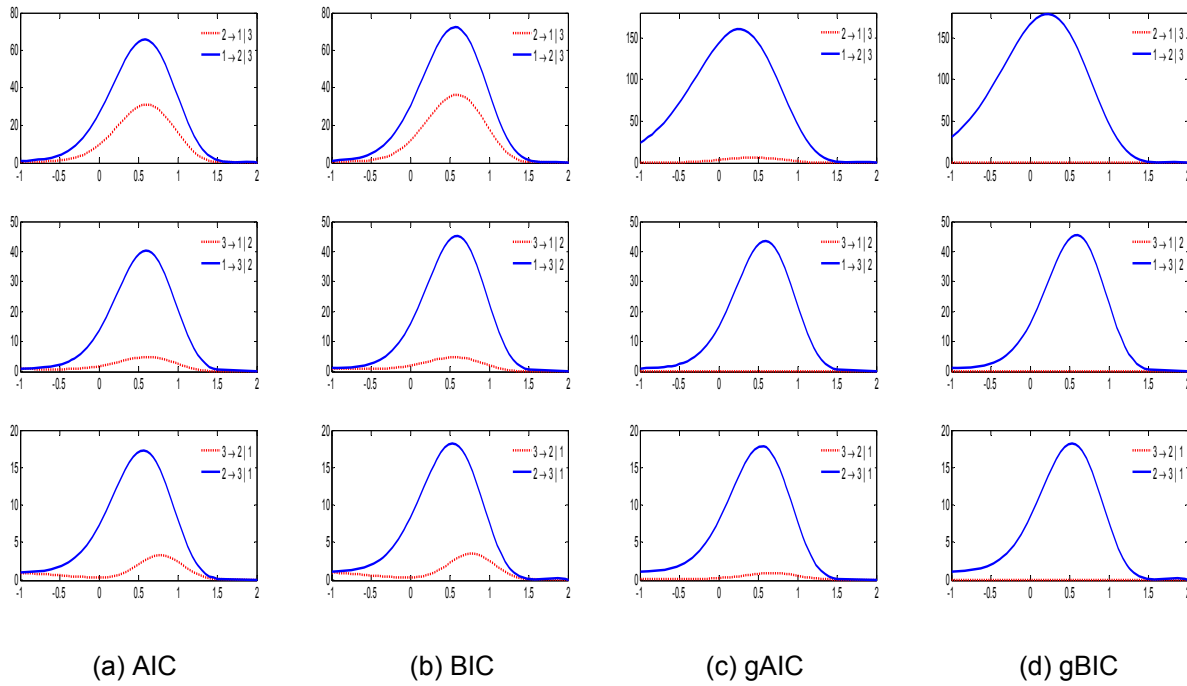


Fig. 5.16 – Results on CTE (bits) with the orders estimated by the four criteria for model 2 (in Eq. (5.1)).
For each criterion, the six available CTE are plotted in one column.

$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$		$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	31.0653 (10.7589)	4.7703 (0.9403)		$j = 1$	-	36.2480 (4.6828)	4.6876 (0.9109)
$j = 2$	65.7285 (12.7757)	-	2.6798 (0.8937)		$j = 2$	72.3068 (5.2741)	-	2.6071 (0.7804)
$j = 3$	40.3670 (9.5429)	17.2043 (2.1058)	-		$j = 3$	45.0429 (4.1473)	18.2034 (1.5060)	-
(a) AIC					(b) BIC			
$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$		$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	5.6485 (15.7183)	0.1499 (2.9156)		$j = 1$	-	0.0405 (1.4068)	0.0000 (0.0000)
$j = 2$	160.2764 (36.0673)	-	0.4539 (1.4085)		$j = 2$	178.7841 (8.9947)	-	0.0000 (0.0000)
$j = 3$	29.7885 (4.3539)	13.8143 (2.1998)	-		$j = 3$	28.1544 (2.4908)	13.2170 (1.2595)	-
(c) gAIC					(d) gBIC			

Tab. 5.12 – Results on CTE at specific r^* for model 2 (in Eq. (5.1)).
The first line indicates the mean and the second line in brackets is the sd.

(a) The orders are estimated by AIC.

(b) The orders are estimated by BIC.

(c) The orders are estimated by gAIC.

(d) The orders are estimated by gBIC.

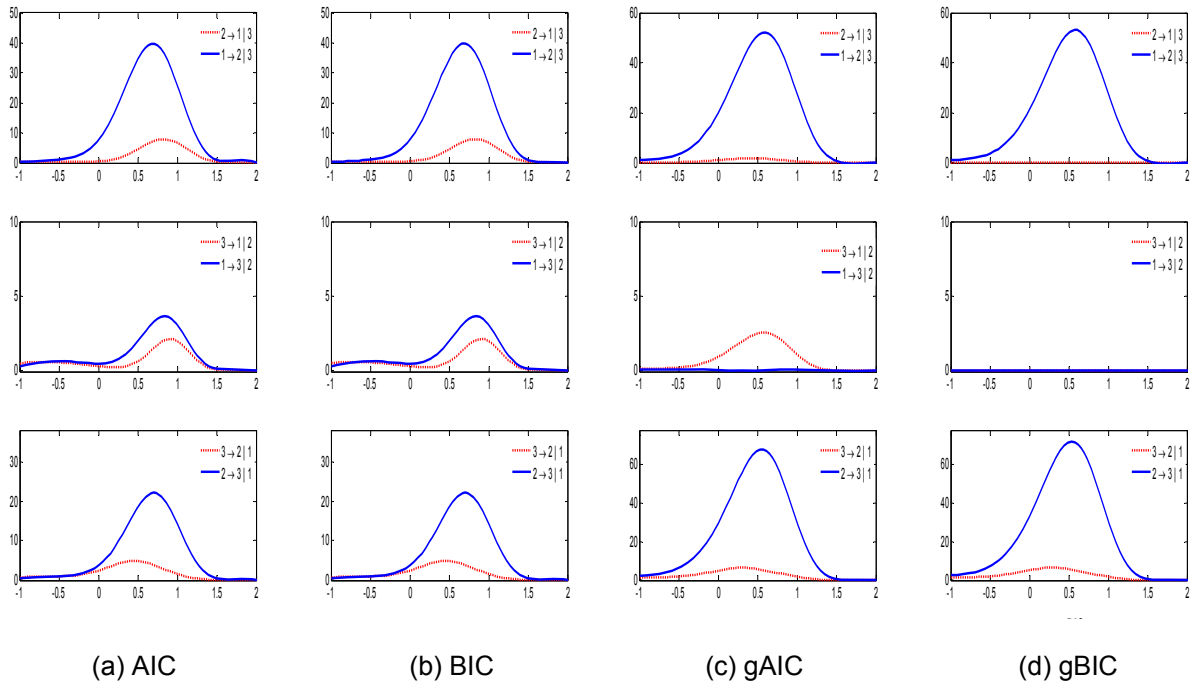


Fig. 5.17 – Results on CTE (bits) with the orders estimated by the four criteria for model 3 (in Eq. (5.1)).
For each criterion, the six available CTE are plotted in one column.

$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$				
$j = 1$	-	7.2187 (1.0464)	1.5102 (0.6710)	$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 2$	39.7103 (1.9062)	-	3.7713 (0.9161)	$j = 1$	-	7.2419 (0.9831)	1.5124 (0.6707)
$j = 3$	3.2753 (0.6965)	22.0989 (1.2321)	-	$j = 2$	39.7665 (1.8312)	-	3.7754 (0.9100)
(a) AIC				$j = 3$	3.2774 (0.6928)	22.1262 (1.1766)	-
				(b) BIC			
$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$				
$j = 1$	-	1.5734 (5.7892)	2.5241 (6.7530)	$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 2$	52.1654 (4.0370)	-	5.0395 (1.5135)	$j = 1$	-	0.0000 (0.0000)	0.0000 (0.0000)
$j = 3$	0.0247 (0.4889)	67.3558 (8.4909)	-	$j = 2$	53.0688 (2.2701)	-	4.9737 (1.4418)
(c) gAIC				$j = 3$	0.0000 (0.0000)	71.0901 (3.6326)	-
				(d) gBIC			

Tab. 5.13 – Results on CTE at specific r^* for model 3 (in Eq. (5.1)).
The first line indicates the mean and the second line in brackets is the sd.

(a) The orders are estimated by AIC.

(b) The orders are estimated by BIC.

(c) The orders are estimated by gAIC.

(d) The orders are estimated by gBIC.

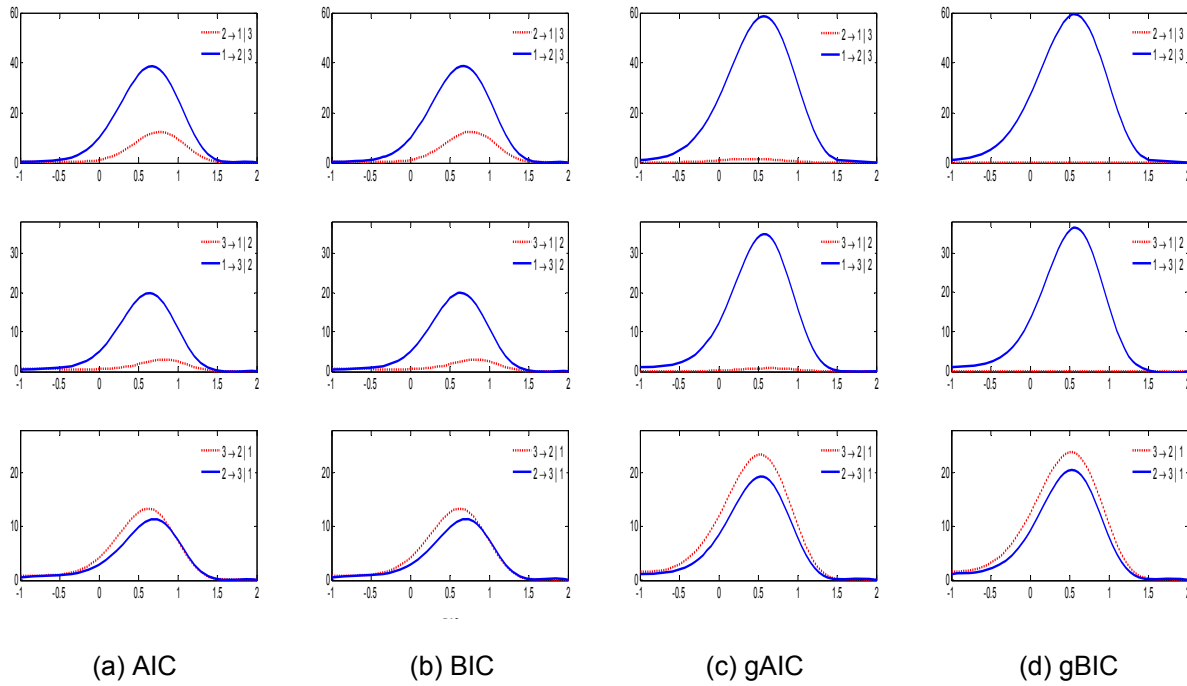


Fig. 5.18 – Results on CTE (bits) with the orders estimated by the four criteria for model 4 (in Eq. (5.1)).
For each criterion, the six available CTE are plotted in one column.

$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$		$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	11.5718 (1.2566)	2.5390 (0.7877)		$j = 1$	-	11.6085 (1.1735)	2.5372 (0.7872)
$j = 2$	38.6048 (2.0097)	-	13.3109 (0.9014)		$j = 2$	38.6711 (1.7018)	-	13.3188 (0.8900)
$j = 3$	19.8607 (1.2835)	11.2320 (0.9979)	-		$j = 3$	19.8859 (1.2401)	11.2633 (0.9303)	-
(a) AIC					(b) BIC			
$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$		$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	1.4502 (8.0930)	0.7460 (4.6511)		$j = 1$	-	0.0405 (1.4068)	0.0000 (0.0000)
$j = 2$	58.6482 (3.7421)	-	23.3890 (2.0638)		$j = 2$	59.4028 (2.4736)	-	23.8463 (1.6062)
$j = 3$	34.7811 (4.5784)	19.3333 (3.6245)	-		$j = 3$	36.3868 (2.6319)	20.5547 (1.8622)	-
(c) gAIC					(d) gBIC			

Tab. 5.14 – Results on CTE at specific r^* for model 4 (in Eq. (5.1)).
The first line indicates the mean and the second line in brackets is the sd.

(a) The orders are estimated by AIC.

(b) The orders are estimated by BIC.

(c) The orders are estimated by gAIC.

(d) The orders are estimated by gBIC.

For each model, the CTE (bits) values are firstly drawn against the corresponding radius r (in the range from $e^{-1.0}$ to $e^{2.0}$ with a step equal to $e^{0.05}$) in log scale in the figures. Subsequently, with the purpose of more visualized analysis, the CTE values at specific r^* are obtained and presented in the tables just after the corresponding figures. For instance, for model 1, the CTE (bits) values with the orders estimated by the four criteria are plotted in Fig. 5.15. Then, the CTE values at specific r^* are displayed in Tab. 5.11. For the remaining 3 models, the homologous CTE results are drawn and shown in Fig. 5.16 and Tab. 5.12 (model 2), Fig. 5.17 and Tab. 5.13 (model 3), and Fig. 5.18 and Tab. 5.14 (model 4) respectively.

From results displayed in Figs. 5.15, 5.16, 5.17, 5.18 and Tabs. 5.11, 5.12, 5.13, 5.14, we conclude that the gBIC criterion is the best criterion (in terms of robustness) to evaluate the model order for estimating TE: (i) $CTE_{i \rightarrow j|k}$ (gBIC) gives the greatest values when the direction of information flow from signal i to signal j conditional to signal k exists, and (ii) $CTE_{i \rightarrow j|k}$ (gBIC) is closer to zero when there is no information flow from signal i to signal j conditional to signal k .

Results on CTE with orders estimated by gBIC presented in Tabs. 5.11, 5.12, 5.13 and 5.14 are summarized in Fig. 5.19. From this figure, we see that CTE is able to detect and discriminate the direct or indirect relations from signal 1 to signal 3 whatever unidirectional or bidirectional relation exists between signals 2 and 3.

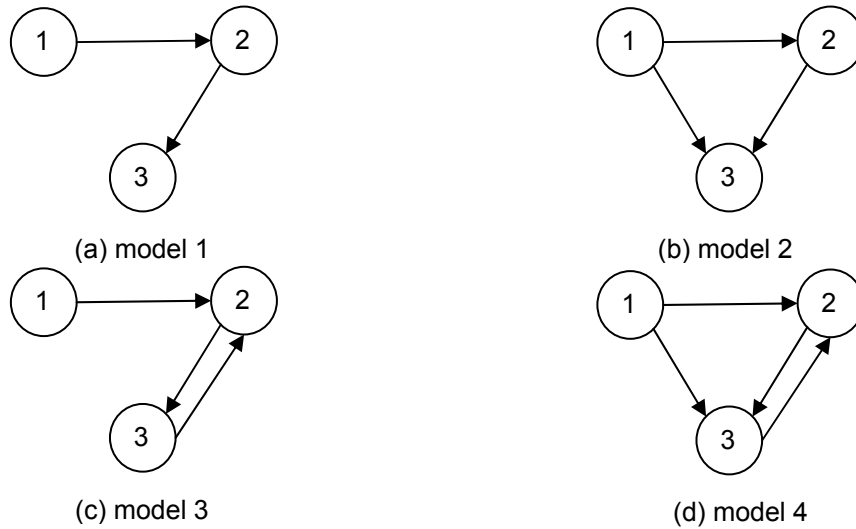


Fig. 5.19 – Estimated graphs from the results on CTE with the orders estimated by gBIC shown in Tabs. 5.11, 5.12, 5.13 and 5.14 for the four linear models.

5.1.3.3. Discussion and conclusion

After comparative analysis of the results in sections 5.1.3.1 and 5.1.3.2, we draw the following conclusions:

(i) Making use of two new criteria of gAIC and gBIC to evaluate the model orders, the performance of TE as well as CTE in detecting causal connectivity is increased in comparison to AIC

and BIC, the gBIC criterion being the most prominent one. In addition, two new "greedy" strategies allow us to reduce drastically the "crippling" computational time.

(ii) When employing gBIC, TE performs well in detecting unidirectional and bidirectional causal relations. However, its pairwise-analysis based structure remains a trouble to distinguish between direct and indirect relations. In order to deal with this issue, the CTE method is proposed to identify and differentiate direct and indirect relations in trivariate systems.

5.2. Second selection of techniques from nonlinear AR models

The stochastic model for the nonlinear signals is governed by the following equations:

$$\begin{cases} x_1(t) = 3.4x_1(t-1)(1-x_1^2(t-1))e^{-x_1^2(t-1)} + w_1(t) \\ x_2(t) = 3.4x_2(t-1)(1-x_2^2(t-1))e^{-x_2^2(t-1)} - 0.5x_1(t-1) + 0.25\sqrt{2}x_2(t-1) - \beta x_3(t-3) + w_2(t) \\ x_3(t) = 3.4x_3(t-1)(1-x_3^2(t-1))e^{-x_3^2(t-1)} - \alpha x_1(t-2) - 0.5x_2(t-2) - 0.25\sqrt{2}x_3(t-2) + w_3(t) \end{cases} \quad (5.3)$$

where $w_j(t)$, $j = 1, 2, 3$, are realizations of independent white noises W_j with zero mean and unit variance. In order to investigate effective connectivity, just like for linear AR models, four nonlinear models including different patterns of causal interactions are considered (see Eq. (5.3) and Fig. 5.1). In Eq. (5.3), the introduction of the parameter α allows to consider two patterns of causal interactions, either direct relations ($\alpha = 0.5$) or indirect relations ($\alpha = 0$). The parameter β ($\beta = 0.5$) is introduced to model bidirectional flow between signals 2 and 3. Therefore, the pair constituted of models 1 and 2 is used to detect and distinguish the direct and indirect relations from signal 1 to signal 3. The other pair of models, models 3 and 4, is also employed to examine direct and indirect relations from signal 1 to signal 3 given a bidirectional relation between signals 2 and 3. Finally, the pair constituted of models 1 and 3 (as well as the pair of models 2 and 4) is regarded for unidirectional and bidirectional relations. In this section, simulations are also carried out 200 times on 2048-point signals and tested to investigate the performance of techniques selected from the previous analysis (section 5.1).

5.2.1. Granger Causality

From section 5.1.1, we found that the WGCI-C (resp. FGCI-C) index behaved better than the WGCI-P (resp. FGCI-P) index in the time (resp. frequency) domain. So, for this family of technique, only results on WGCI-C and FGCI-C indices are presented and analyzed in the following sections.

5.2.1.1. WGCI-C

The results on WGCI-C for the four nonlinear models (in Eq. (5.3)) are summarized in Tab. 5.15. The first line indicates the mean value, and the second line in brackets is the sd. The corresponding synthesized schemas are shown in Fig. 5.20.

Following the same analysis as for the linear AR models, and examining the pair of models 1 and 2 in Tab. 5.15, we find that the WGCI-C index allows us to identify the direct causality relation. It distinguishes properly the two patterns of causal interactions (indirect ($WGCI_{1 \rightarrow 3}\text{-C} = 0.0010$ in Tab.

5.15a) vs direct ($WGCI_{1 \rightarrow 3} - C = 0.1900$ in Tab. 5.15b)). The same conclusion holds comparing models 3 and 4 ($WGCI_{1 \rightarrow 3} - C = 0.0011$ in Tab. 5.15c vs $WGCI_{1 \rightarrow 3} - C = 0.2009$ in Tab. 5.15d). The propagation graphs in Fig. 5.20 are in total concordance with those of Fig. 5.1, confirming that the $WGCI - C$ index succeeds in differentiating the direct or indirect causal interrelations from signal 1 to signal 3 whatever unidirectional or bidirectional relations exist between these two signals.

$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	0.0010 (0.0004)	0.0009 (0.0004)
$j = 2$	0.2274 (0.0196)	-	0.0010 (0.0003)
$j = 3$	0.0010 (0.0003)	0.2911 (0.0218)	-
(a) model 1			
$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	0.0009 (0.0003)	0.0010 (0.0003)
$j = 2$	0.2294 (0.0195)	-	0.0010 (0.0004)
$j = 3$	0.1900 (0.0187)	0.2986 (0.0209)	-
(b) model 2			
$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	0.0013 (0.0005)	0.0012 (0.0004)
$j = 2$	0.2325 (0.0192)	-	0.3333 (0.0228)
$j = 3$	0.0011 (0.0004)	0.3093 (0.0226)	-
(c) model 3			
$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	0.0011 (0.0004)	0.0012 (0.0004)
$j = 2$	0.2313 (0.0211)	-	0.3905 (0.0260)
$j = 3$	0.2009 (0.0185)	0.3234 (0.0237)	-
(d) model 4			

Tab. 5.15 – Results on $WGCI - C$ for the four nonlinear models derived from Eq. (5.3).
The first line indicates the mean and the second line in brackets is the sd.

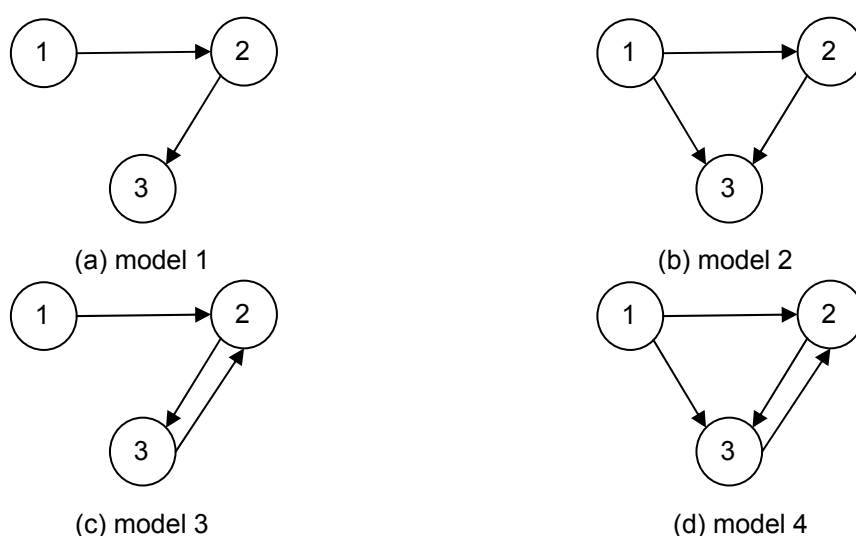


Fig. 5.20 – Estimated graphs from the results on $WGCI - C$ shown in Tab. 5.15 for the four nonlinear models.

5.2.1.2. FGCI-C

Results on the means values of the FGCI-C index at each frequency f for the four nonlinear models (in Eq. (5.3)) are drawn in Fig. 5.21, and the corresponding propagation graphs in Fig. 5.22.

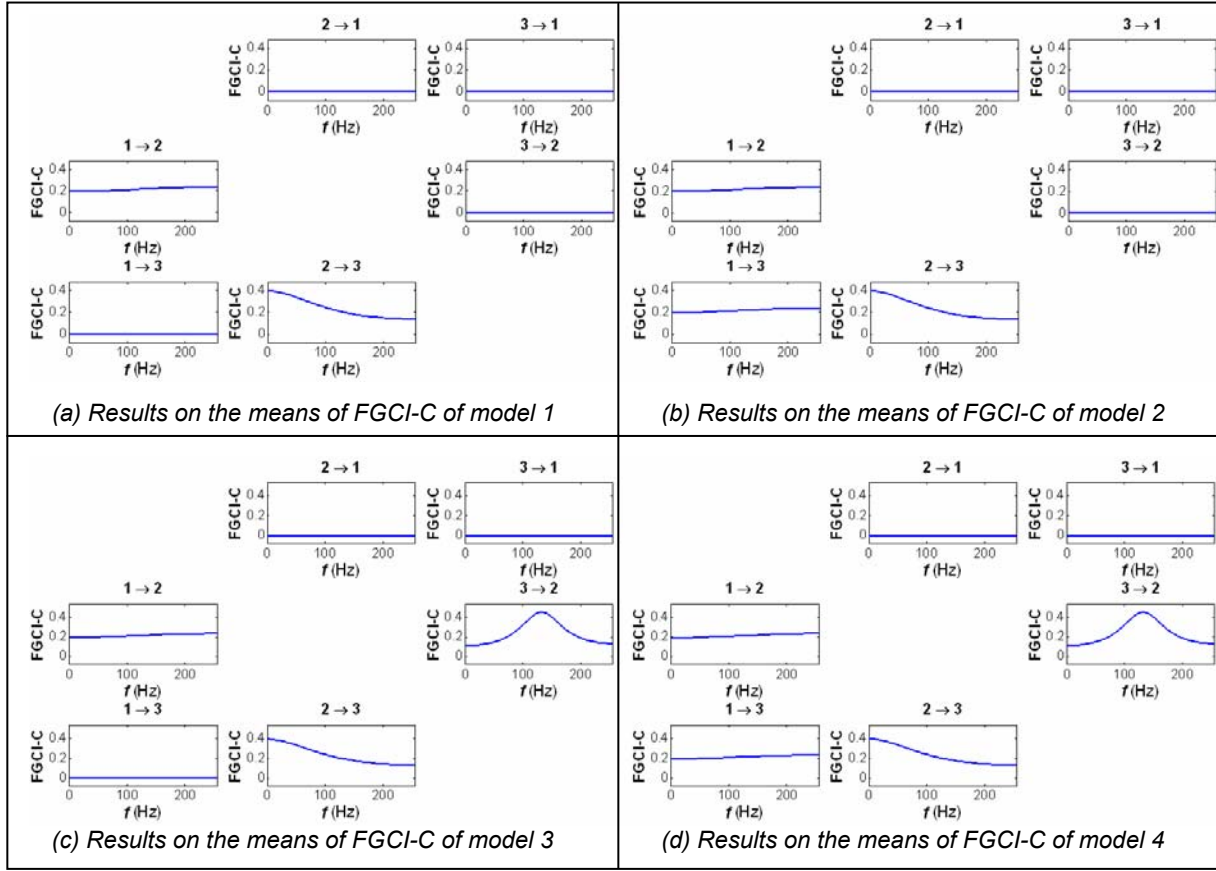


Fig. 5.21 – Mean values of the FGCI-C index at each frequency f for the four nonlinear models derived from Eq. (5.3).

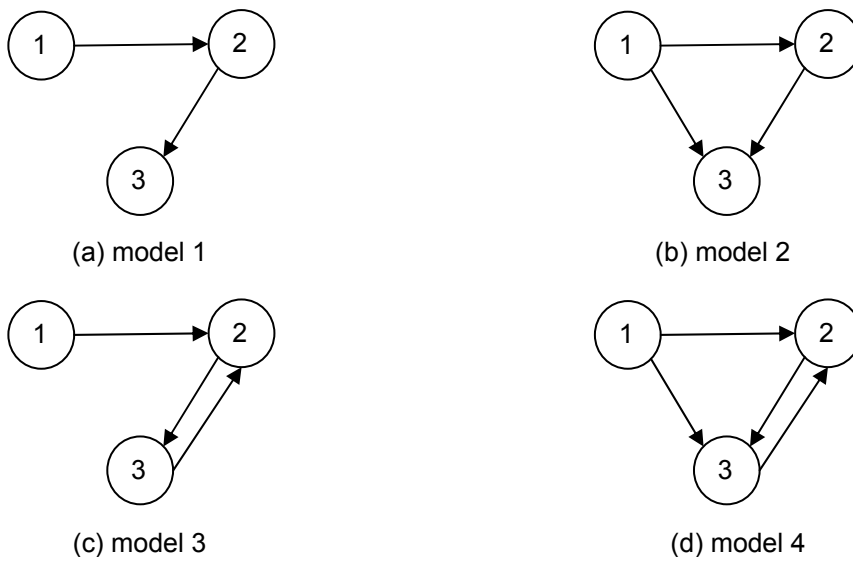


Fig. 5.22 – Estimated graphs from the results on FGCI-C shown in Fig. 5.21 for the four nonlinear models of Eq. (5.3).

As analyzed in linear models, the $FGCI_{1 \rightarrow 3}-C$ index is also able to differentiate the direct or indirect causal relations from signal 1 to signal 3 whatever unidirectional or bidirectional relations exist between these two signals (see models 1 and 2, or models 3 and 4 in Figs. 5.21 and 5.22).

5.2.1.3. Discussion and conclusion

As analyzed for linear models, both $WGCI-C$ and $FGCI-C$ indices succeed in identifying and differentiating various patterns of causal relations in the time or frequency domain respectively in this case of nonlinear models.

5.2.2. Phase slope index and causality index

From section 5.1.2, we know that the $PSI-OC$ and $CI-PC$ indices can only deal with systems involving unidirectional relations. Therefore, in the following, only $CI-DCOH$, $CI-DTF$ and $CI-PDC$ indices are retained and tested on the pair of models 3 and 4 (in Eq. (5.3)) to investigate their performance in detecting different types of relations.

5.2.2.1. $CI-DCOH$ and $CI-DTF$ vs $CI-PDC$

Numerical results on $CI-DCOH$, $CI-DTF$ and $CI-PDC$ for models 3 and 4 are summarized and presented in Tab. 5.16 and the corresponding propagation schemas are shown in Fig. 5.23.

On the one hand, the $CI-DCOH$ index allows for pointing out unidirectional and bidirectional relations (see Tabs. 5.16a and 5.16b, either model 3 or model 4), and, on the other hand, when two different interactions exist from signal 1 to signal 3, *i.e.* indirect (model 3) and direct (model 4, signal 1 connects to signal 3 via two distinct pathways) relations, $CI_{13}-DCOH$ displays systematically important values and, consequently, fails in discriminating direct and indirect relations.

Like $CI-DCOH$, $CI-DTF$ is successful in pointing out unidirectional and bidirectional relations but fails in distinguishing direct and indirect relations from signal 1 to signal 3 in models 3 and 4 (see Tabs. 5.16c and 5.16d).

Like $CI-DCOH$ and $CI-DTF$, from Tabs. 5.16e and 5.16f the $CI-PDC$ index succeeds in detecting unidirectional and bidirectional relations for the two models. Moreover, when there is only indirect flow from signal 1 to signal 3 (relation mediated by signal 2 in model 3), $CI_{13}-PDC$ remains close to zero (Tab. 5.16e), while it increases significantly (Tab. 5.16f) when there is a direct relation from signal 1 to signal 3 (model 4). Thus, $CI-PDC$ clearly contrasts with $CI-DCOH$ and $CI-DTF$, since these two quantities display a non negligible value when tested from signal 1 to signal 3, whatever the value of α . Therefore, $CI-PDC$ answers the question of direct and indirect connections between pairs of signals since it differentiates these kinds of connection.

Furthermore, even if $CI-PDC$ and $CI-DTF$ cannot reveal the delays between signals in the bidirectional case, one can at least infer the value of the delays' ratio between signals 2 and 3 (a 2-time delay from signal 2 to signal 3 and a 3-time delay from signal 3 to signal 2). As a matter of fact, the ratio $CI_{32}-PDC/CI_{23}-PDC$ (as well as the ratio $CI_{32}-DTF/CI_{23}-DTF$) is always around 3/2

regardless of the model, whereas the ratio $CI_{32}\text{-DCOH}/CI_{23}\text{-DCOH}$ varies much more (around 1.14 for model 3 and 1.31 for model 4).

The conclusions drawn from Table 5.16 can be visualized through the propagation graphs shown in Fig. 5.23.

$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$		$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	0.0319 (0.0255)	0.0306 (0.0229)		$j = 1$	-	0.0346 (0.0254)	0.0279 (0.0216)
$j = 2$	0.5527 (0.0607)	-	1.7505 (0.1221)		$j = 2$	1.1604 (0.1474)	-	2.0285 (0.1306)
$j = 3$	0.4925 (0.0886)	1.5386 (0.0916)	-		$j = 3$	1.5282 (0.1171)	1.5436 (0.0971)	-
(a) CI-DCOH on model 3					(b) CI-DCOH on model 4			
$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$		$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	0.0114 (0.0079)	0.0139 (0.0091)		$j = 1$	-	0.0103 (0.0072)	0.0123 (0.0079)
$j = 2$	0.5269 (0.0418)	-	1.7788 (0.1144)		$j = 2$	1.1683 (0.0714)	-	1.7212 (0.0952)
$j = 3$	0.4698 (0.0812)	1.2555 (0.0696)	-		$j = 3$	1.4817 (0.0986)	1.1098 (0.0628)	-
(c) CI-DTF on model 3					(d) CI-DTF on model 4			
$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$		$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	0.0063 (0.0059)	0.0056 (0.0050)		$j = 1$	-	0.0056 (0.0050)	0.0046 (0.0040)
$j = 2$	0.6342 (0.0425)	-	1.9988 (0.1087)		$j = 2$	0.5250 (0.0399)	-	1.9999 (0.1067)
$j = 3$	0.0075 (0.0062)	1.3085 (0.0723)	-		$j = 3$	1.0498 (0.0829)	1.3109 (0.0713)	-
(e) CI-PDC on model 3					(f) CI-PDC on model 4			

Tab. 5.16 – Results on CI-DCOH, CI-DTF and CI-PDC for nonlinear models 3 and 4.

The first line indicates the mean and the second line in brackets is the sd.

5.2.2.2. Discussion and conclusion

Whatever the simulated AR model (linear or nonlinear), the CI-PDC index is able to identify (and so discriminate) direct and indirect relations as well as unidirectional and bidirectional relations whereas the CI-DCOH and CI-DTF indices only succeed in the latter case.

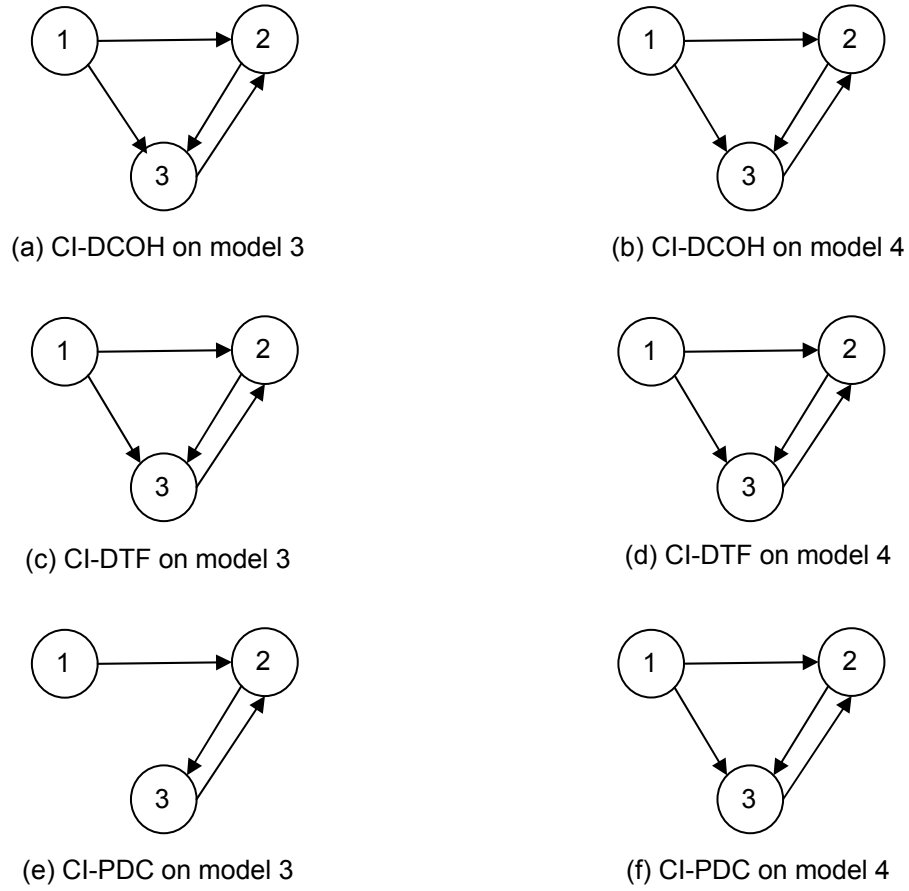


Fig. 5.23 – Estimated graphs from the results on CI-DCOH, CI-DTF and CI-PDC shown in Tab. 5.16 for nonlinear models 3 and 4.

5.2.3. Transfer entropy and conditional transfer entropy

As discussed in section 5.1.3, the criterion to evaluate the model order plays a prominent part in the estimation of TE and CTE indices. According to the previous results, we choose to present in this section only performance of TE and CTE when the orders are estimated using the gBIC criterion including the two respective "greedy" strategies for the four nonlinear AR models.

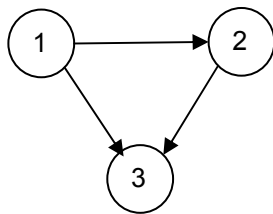
5.2.3.1. TE

For the four nonlinear models, the numerical values of TE are summarized in Tab. 5.17 and the corresponding graphs are plotted in Fig. 5.24.

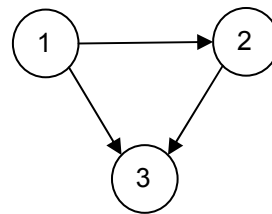
From Tab. 5.17 and the corresponding schemas in Fig. 5.24, it comes out that the TE measure detects unidirectional and bidirectional relations but it fails in distinguishing between direct and indirect relations.

$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$		$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	0.0130 (0.1943)	0.0000 (0.0000)		$j = 1$	-	0.0000 (0.0000)	0.0000 (0.0000)
$j = 2$	40.5745 (3.7144)	-	-0.0012 (0.0118)		$j = 2$	40.6195 (3.8792)	-	0.0104 (0.1471)
$j = 3$	4.4968 (1.4566)	29.1930 (2.4397)	-		$j = 3$	20.8780 (2.1308)	27.4902 (2.0867)	-
(a) model 1					(b) model 2			
$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$		$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	0.0043 (0.0603)	0.0000 (0.0000)		$j = 1$	-	0.0000 (0.0000)	0.0129 (0.1824)
$j = 2$	13.2792 (1.6990)	-	30.0165 (2.7645)		$j = 2$	9.4508 (1.1662)	-	35.3308 (2.8621)
$j = 3$	2.7491 (0.9552)	36.1319 (2.4667)	-		$j = 3$	12.4236 (1.2528)	30.4616 (4.8442)	-
(c) model 3					(d) model 4			

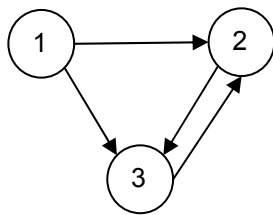
Tab. 5.17 – Results on TE for the four nonlinear models derived from Eq. (5.3).
The first line indicates the mean and the second line in brackets is the sd.



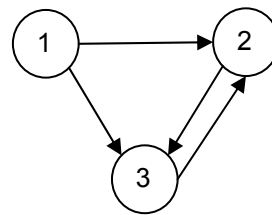
(a) model 1



(b) model 2



(c) model 3



(d) model 4

Fig. 5.24 – Estimated graphs from the results on TE shown in Tab. 5.17 for the four nonlinear models.

5.2.3.2. CTE

Results on CTE including the gBIC technique are shown in Tab. 5.18 for the four nonlinear models and the corresponding propagation graphs are drawn in Fig. 5.25.

From Tab. 5.18 and Fig. 5.25, the CTE index performs well to detect and differentiate different patterns of causal interactions existing in the four nonlinear models.

$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	0.0106 (0.1767)	0.0000 (0.0000)
$j = 2$	40.4323 (3.9125)	-	0.0009 (0.0157)
$j = 3$	0.0000 (0.0000)	27.4448 (2.4336)	-
(a) model 1			
$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	0.0003 (0.0727)	0.0000 (0.0000)
$j = 2$	16.3979 (1.5283)	-	25.2882 (2.1316)
$j = 3$	0.0088 (0.0922)	36.0283 (2.5831)	-
(c) model 3			
$i \rightarrow j$	$i = 1$	$i = 2$	$i = 3$
$j = 1$	-	0.0000 (0.0000)	0.0101 (0.1422)
$j = 2$	15.5102 (1.6573)	-	29.5944 (2.4082)
$j = 3$	12.4832 (1.3653)	21.8340 (1.7435)	-
(d) model 4			

Tab. 5.18 – Results on CTE for the four nonlinear models derived from Eq. (5.3).
The first line indicates the mean and the second line in brackets is the sd.

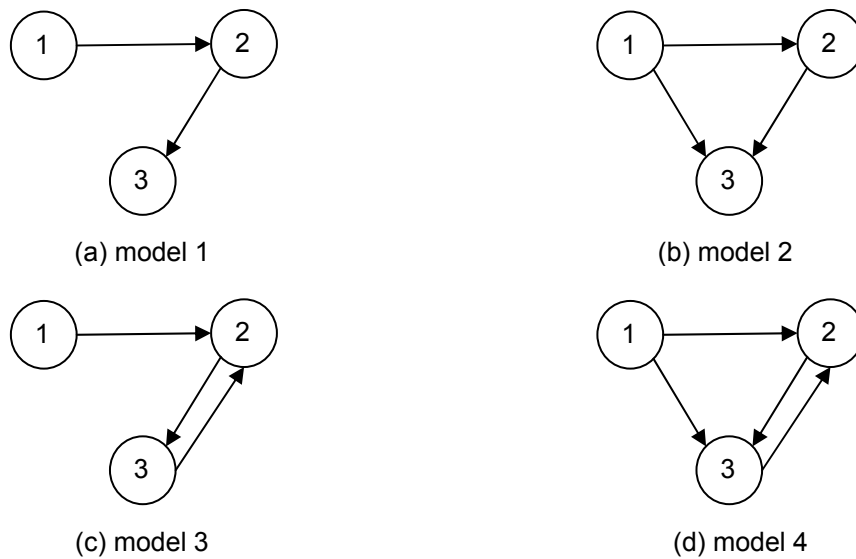


Fig. 5.25 – Estimated graphs from the results on CTE with the orders estimated by gBIC shown in Tab. 5.18 for the four nonlinear models.

5.2.3.3. Discussion and conclusion

Whatever linear or nonlinear models, the TE index is able to detect and distinguish the unidirectional or bidirectional relations. Besides, the CTE index further succeeds in identifying and discriminating direct and indirect relations.

5.3. Validation phase

For the physiological models and real signals, the situations are more complicated. For instance, the difference between "weak" and "no" relation can be difficult to judge on the mean values and standard deviations due to outliers. Therefore, a statistical analysis seems more appropriate. In this validation phase, we decide to test only three methods according to the results obtained in the two previous sections. So, box plots are chosen to display statistical results on WGCI-C, CI-PDC and CTE for the same type of connectivity (model 4) not only on the physiological model but also on the linear and nonlinear models (model 4, too). For each configuration, the results for all techniques are plotted in a same figure. In this figure, the six available indices (given a set of 3 signals) are plotted for each method. In the following tables, a pair (i, j) corresponds to the information flow from signal i to signal j . The left box plot is relative to \bar{H}_0 hypothesis, and the right box plot to H_0 hypothesis (null hypothesis).

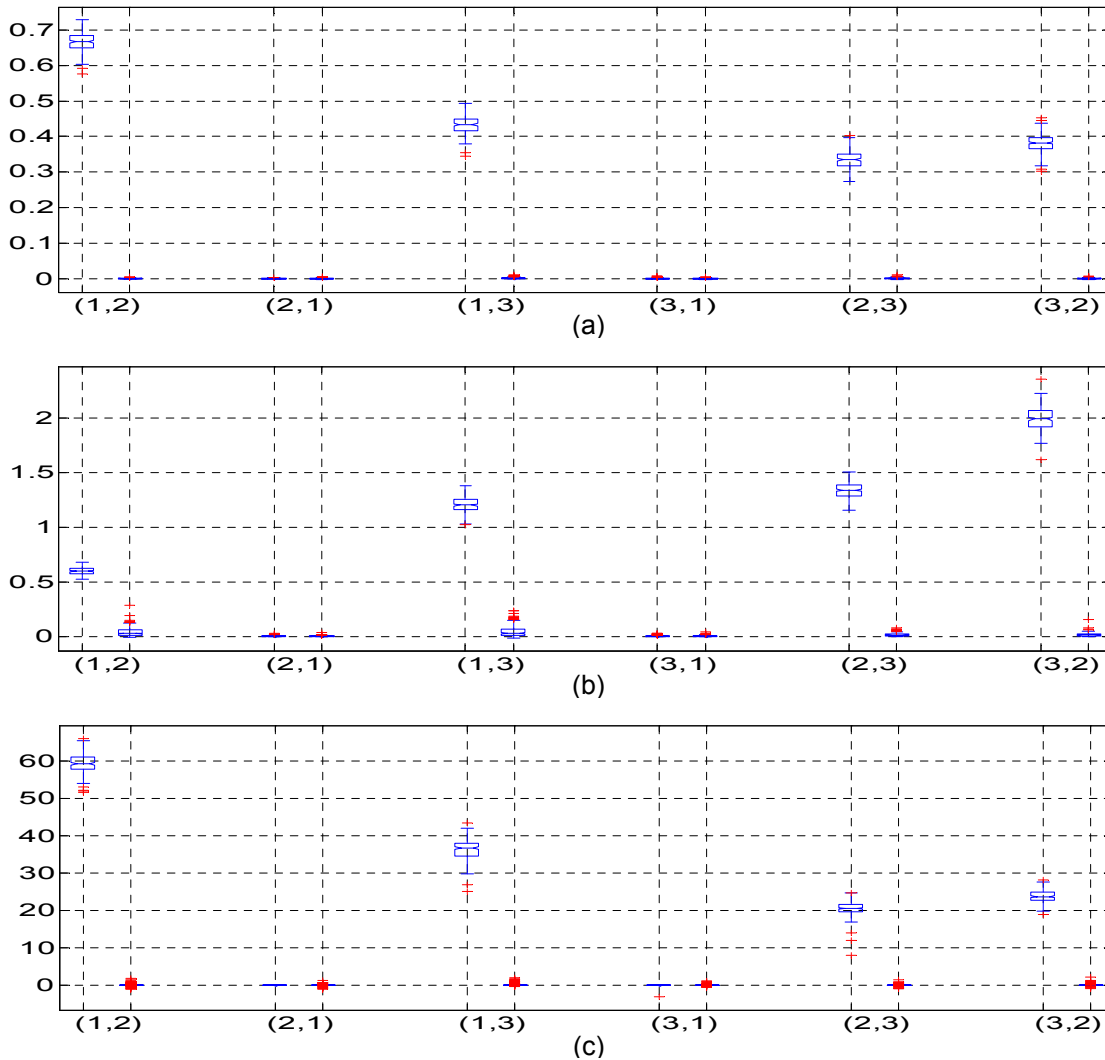


Fig. 5.26 – Box plot of indices (a) WGCI-C, (b) CI-PDC and (c) CTE for the linear model 4
x-axis: (i, j) presents the information flow from signal i to signal j ,
under H_0 hypothesis (right) and under \bar{H}_0 hypothesis (left)
y-axis: results under H_0 hypothesis (right) and under \bar{H}_0 hypothesis (left) for the pair (i, j) .

5.3.1. Results on the linear model (model 4)

Results on WGCi-C, CI-PDC and CTE are plotted in Fig. 5.26 and show that the three techniques detect and differentiate properly all types of causal relations as already seen in Fig. 5.1(d). As a matter of fact, results under H_0 and \bar{H}_0 are well distinguished for the following links: $1 \rightarrow 2$, $1 \rightarrow 3$ and $2 \rightleftharpoons 3$. Moreover, for the CI-PDC index, it somewhat "quantifies" the delays: if we compare the links $3 \rightarrow 2$ and $2 \rightarrow 3$, we note a ratio of 3/2 corresponding to the delays chosen in Eq. (5.1).

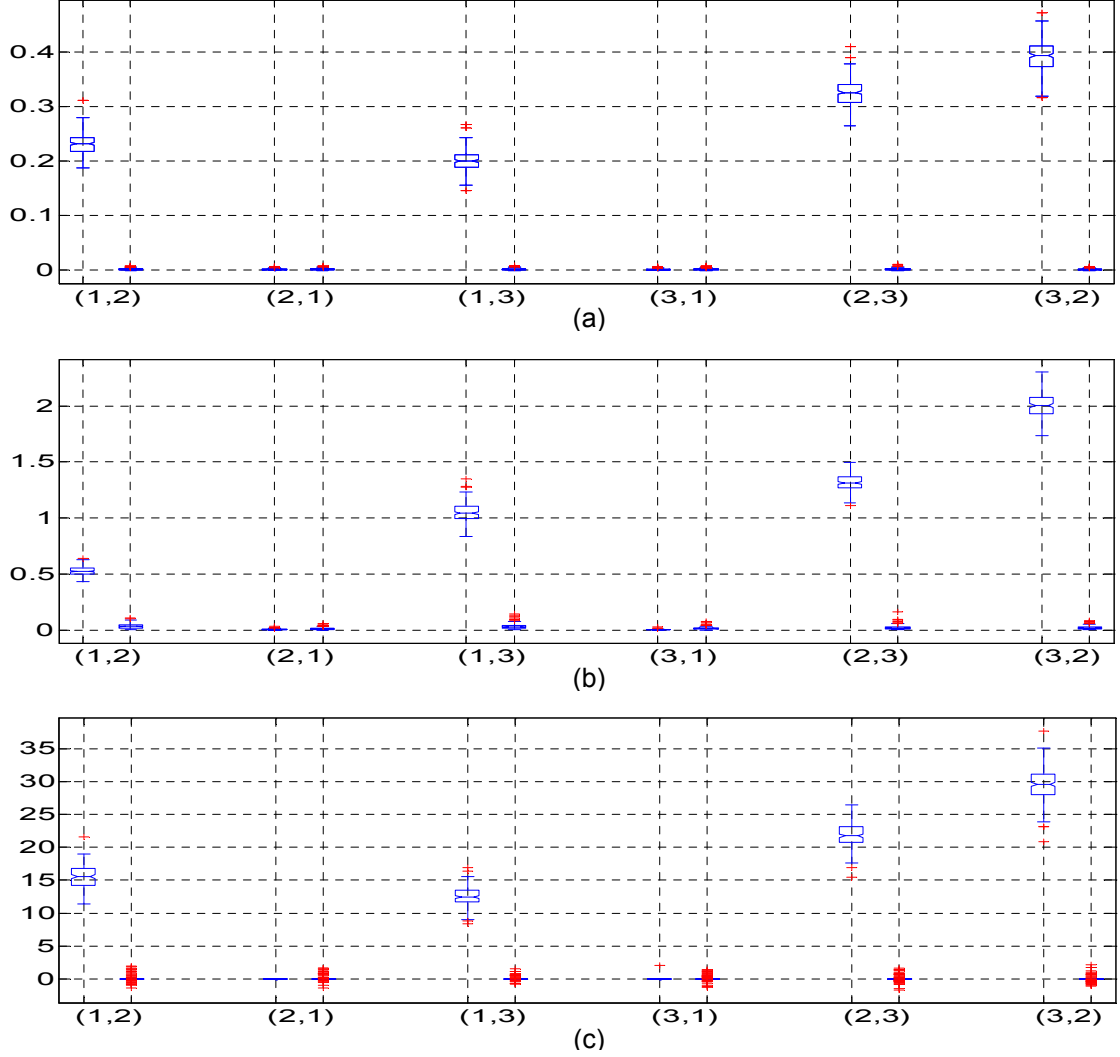


Fig. 5.27 – Box plot of indices (a) WGCi-C, (b) CI-PDC and (c) CTE for the nonlinear model 4
*x-axis: (i, j) presents the information flow from signal i to signal j ,
under H_0 hypothesis (right) and under \bar{H}_0 hypothesis (left)*
y-axis: results under H_0 hypothesis (right) and under \bar{H}_0 hypothesis (left) for the pair (i, j) .

5.3.2. Results on the nonlinear model (model 4)

Results on WGCi-C, CI-PDC and CTE are plotted in Fig. 5.27 and allow to draw the same conclusion as in the linear case. There is no ambiguity on the true relations between signals and the

underlying model can be perfectly identified. Results under H_0 and \bar{H}_0 are only overlapped for the two directions $2 \rightarrow 1$ and $3 \rightarrow 1$ confirming the independence in these two cases (no causal influence).

5.3.3. Results on the physiological model (model 4)

In this section, WGCI-C, CI-PDC and CTE are tested on the physiological model with the following parameters $K^{12} = K^{13} = K^{23} = K^{32} = 1500$. Using the model 4 described in chapter 3, we generated 800-s signals with sampling frequency 512 Hz. Then, these signals were divided into 200 non overlapped segments and we computed each index on each segment. The corresponding results are drawn in Fig. 5.28.

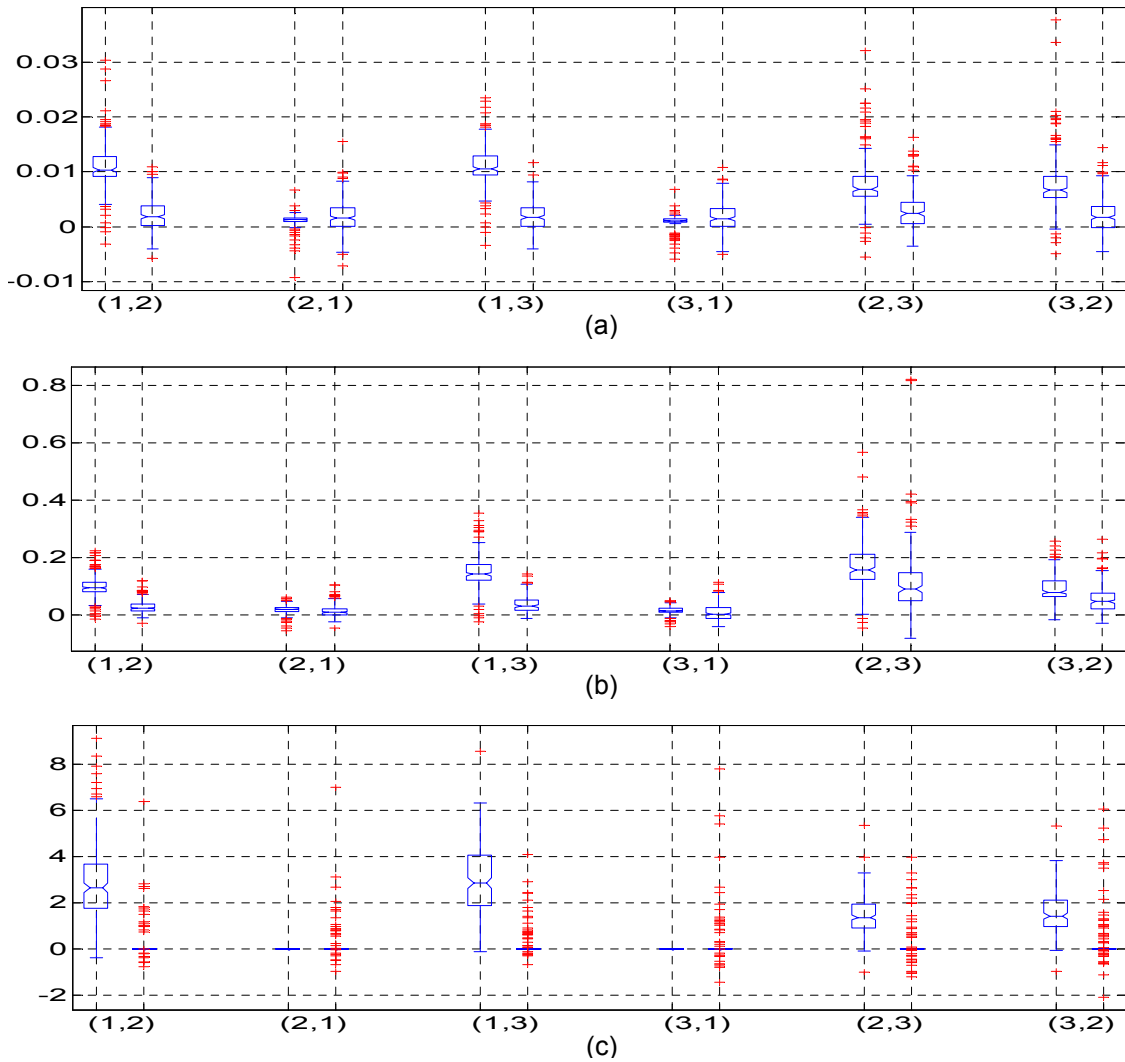
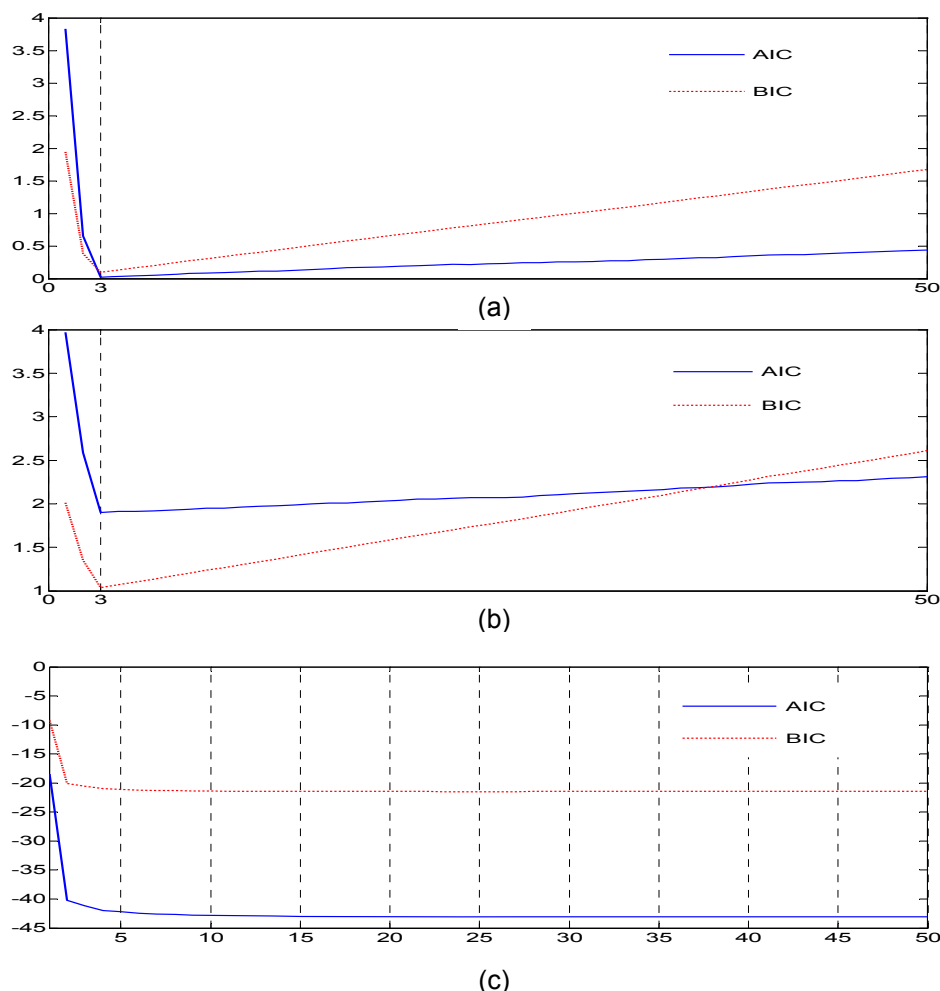


Fig. 5.28 – Box plot of indices (a) WGCI-C, (b) CI-PDC and (c) CTE for physiological model 4
x axis: (i, j) presents the information flow from signal i to signal j ,
under H_0 hypothesis (right) and under \bar{H}_0 hypothesis (left)
y axis: results under H_0 hypothesis (right) and under \bar{H}_0 hypothesis (left) for the pair (i, j) .

For the three techniques, a criterion is needed to estimate the orders of the models. For the linear and nonlinear models, the optimal orders are easily obtained (see Fig. 5.29(a) and Fig. 5.29(b):

the optimal order is obviously equal to 3)). However, for the physiological model, the situation is more complicated and the optimal order can be over-estimated as can be seen from Fig. 5.29(c): the two curves decrease rapidly from the starting point and then they remain quite constant and the optimal order can be large. Hence, to take into account the computation time, the maximal order is set to 6. As in sections 5.1.3.2 and 5.2.3.2, CTE is tested here using gBIC.



*Fig. 5.29 – Estimation of the order using AIC and BIC
(a) linear model 4, (b) nonlinear model 4, (c) physiological model 4
x-axis: values of the model order*

y-axis: values of the criteria (AIC (blue solid line) and BIC (red dashed line)) for each order.

From Fig. 5.28, we note that H_0 and \bar{H}_0 hypotheses can be distinguished for all techniques for the unidirectional relations $1 \rightarrow 2$ and $1 \rightarrow 3$. Results are not so obvious on the bidirectional relation $2 \rightleftharpoons 3$. Getting rid of outliers, CTE seems relevant, thanks to the optimization of the order. Consequently, we can expect some comparable results with WGCI-C in the same configuration (which has not been considered until now. As for the CI-PDC index, it seems to be less robust. Since Fig. 5.28 reveals a dispersion of the values, whatever the method, we represent on Fig. 5.30 the results using ROC (Receiver Operating Characteristic) curves.

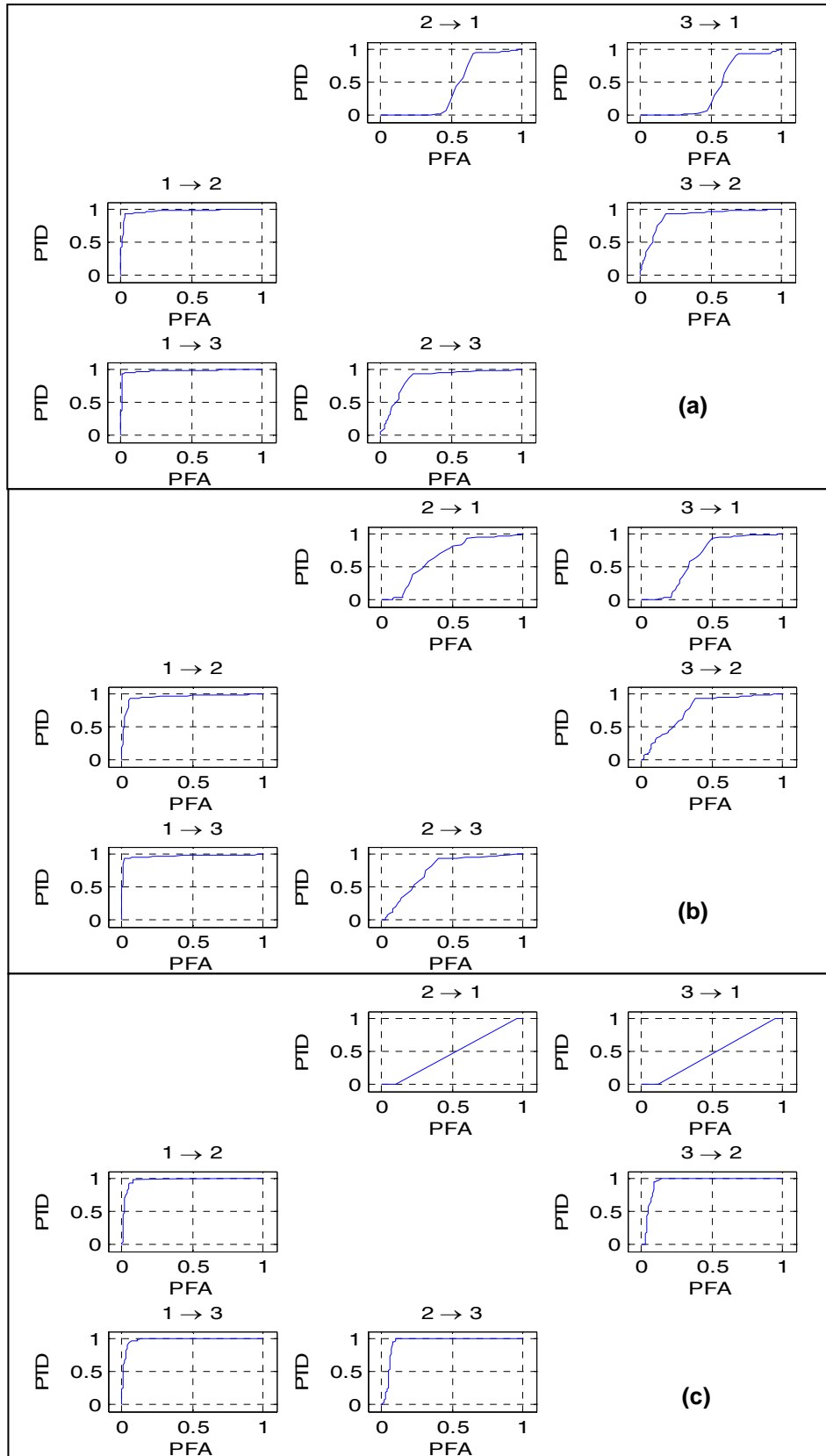


Fig. 5.30 – ROC curve on (a) WGCI-C, (b) CI-PDC and (c) CTE for the physiological model 4.
x-axis: probability of false alarm (PFA)
y-axis: probability of true detection (PTD).

From Fig. 5.30, for the unidirectional relations, *i.e.* $1 \rightarrow 2$ and $1 \rightarrow 3$, results obtained by the three techniques are quite perfect. For the bilateral relation $2 \rightleftharpoons 3$, CTE is successful, while CI-PDC confirms its weakness. Note that the curves corresponding to the non relations $2 \rightarrow 1$ and $3 \rightarrow 1$ are close to the diagonal, which is normal since the two experimental distributions correspond in this case to the same hypothesis, H_0 , and, therefore, overlap.

Finally, as an example, we represent the ROC curves for a given relation, namely $3 \rightarrow 2$, in the case of a nonlinear model and a physiological one (for the same type of connectivity), to appreciate and highlight the performance of the three techniques in both cases. Obviously, these three techniques are all relevant in the nonlinear situation. As for the physiological model, CTE still behaves well while CI-PDC has more trouble.

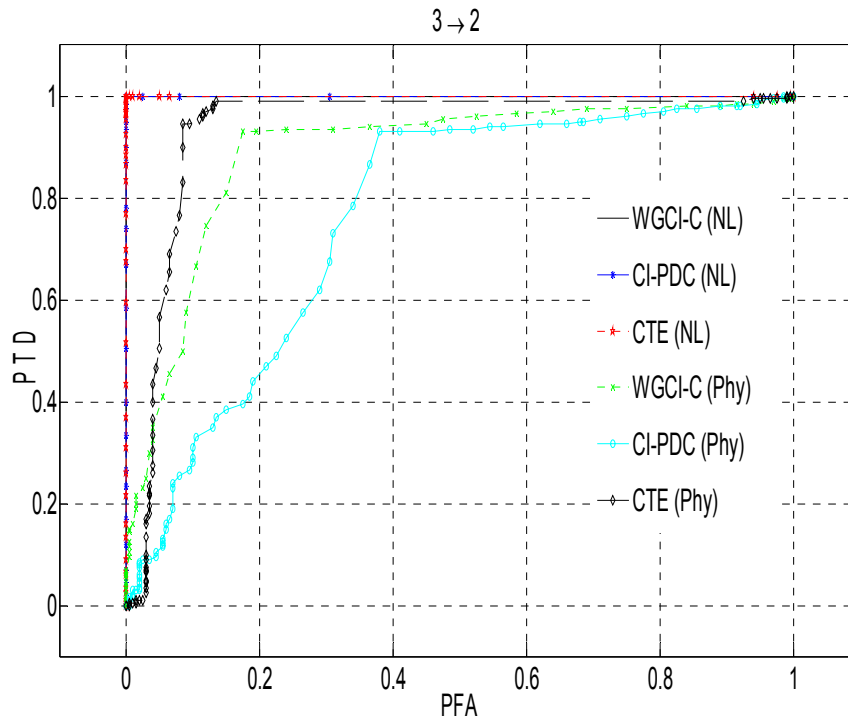


Fig. 5.31 – ROC curve on (a) WGCi-C, (b) CI-PDC and (c) CTE for the relation from 3 to 2 for the nonlinear and physiological models
 NL: nonlinear model 4, Phy: physiological model 4
 x-axis: probability of false alarm (PFA)
 y-axis: probability of true detection (PTD).

5.4. Tests on real signals

As discussed in chapter 3, we analyzed three real signals recorded during ictal onset. This phase is characterized by a peculiar pattern of fast activity that originates from the entorhinal/hippocampal region and secondarily propagates to the perirhinal cortex [Uva 2005]. These signals are tested with WGCi-C, CI-PDC, and CTE techniques. Signal 1 is recorded in the hippocampus, signal 2 in the medial entorhinal cortex and signal 3 in the perirhinal cortex. The time duration of the fast onset activity phase we considered was 10 seconds. To detect the relations at

different time instants, a 4-s sliding window with a step of 0.125 s (a total of 49 windows) and an optimal order of 8 were chosen for the estimation of all indices (see Fig. 5.32).

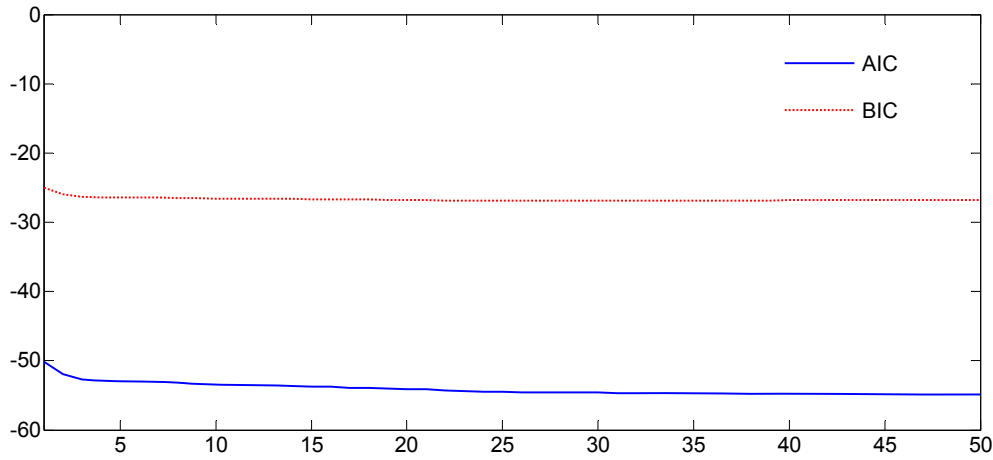


Fig. 5.32 – Estimation of the order using AIC and BIC for the three recorded signals on a 4-s window.

In Fig. 5.33, we represented the evolution of the different indices - WGCI-C, CI-PDC, and CTE - over time during the analyzed period corresponding to this Fast Onset Activity (FOA) phase. The significance of the values of indices revealing the influence of signal i on signal j , and *vice versa*, and their variations can be established using the box plots represented respectively at the bottom-left for the relation from i to j , and, at the bottom-right for the relation from j to i . These boxes have been obtained with the method of random phase to simulate H_0 hypothesis. More precisely, we considered the whole duration, *i.e.* 10 seconds of observation, and we constructed 4 windows of 4 seconds each, with an overlap of 2 seconds. For each window, we performed 50 realizations randomizing the phase so as to get 200 values (4 x 50) of the index under H_0 merged in these box plots.

The results obtained by the three techniques express the same trend. Signals 1 and 2 seem to influence each other on the whole sequence even if there is a moderate decrease in the influence from signal 1 to signal 2 (especially with CI-PDC and WGCI-C methods) on the latter part of the interval. Regarding the relations linking signals 1 and 3, they invert around the middle of the sequence. The relation from signal 1 to signal 3 is strong at the beginning (same order of magnitude as that of the relation from signal 1 to signal 2) and then decreases to non significant values on the latter part of the interval whatever the method. For all methods, regarding the relation from signal 3 to signal 1, it is absent in the former part of the interval and rises slightly on the latter part. As for the pair of signals (2, 3), a relation from signal 2 to signal 3 is predominant for all methods, mainly on the latter half of the global analysis window. No influence from signal 3 to signal 2 is detected over time. These conclusions are visually synthesized in Fig. 5.34 on the two parts of the interval. These results seem globally consistent but they bring to light the sensitivity of the indices and the possible reversal of the relations on short time intervals. They must be taken with care since they come from only one recording whereas experiments reported in [Uva 2005] were performed in more than thirty isolated guinea pig brains.

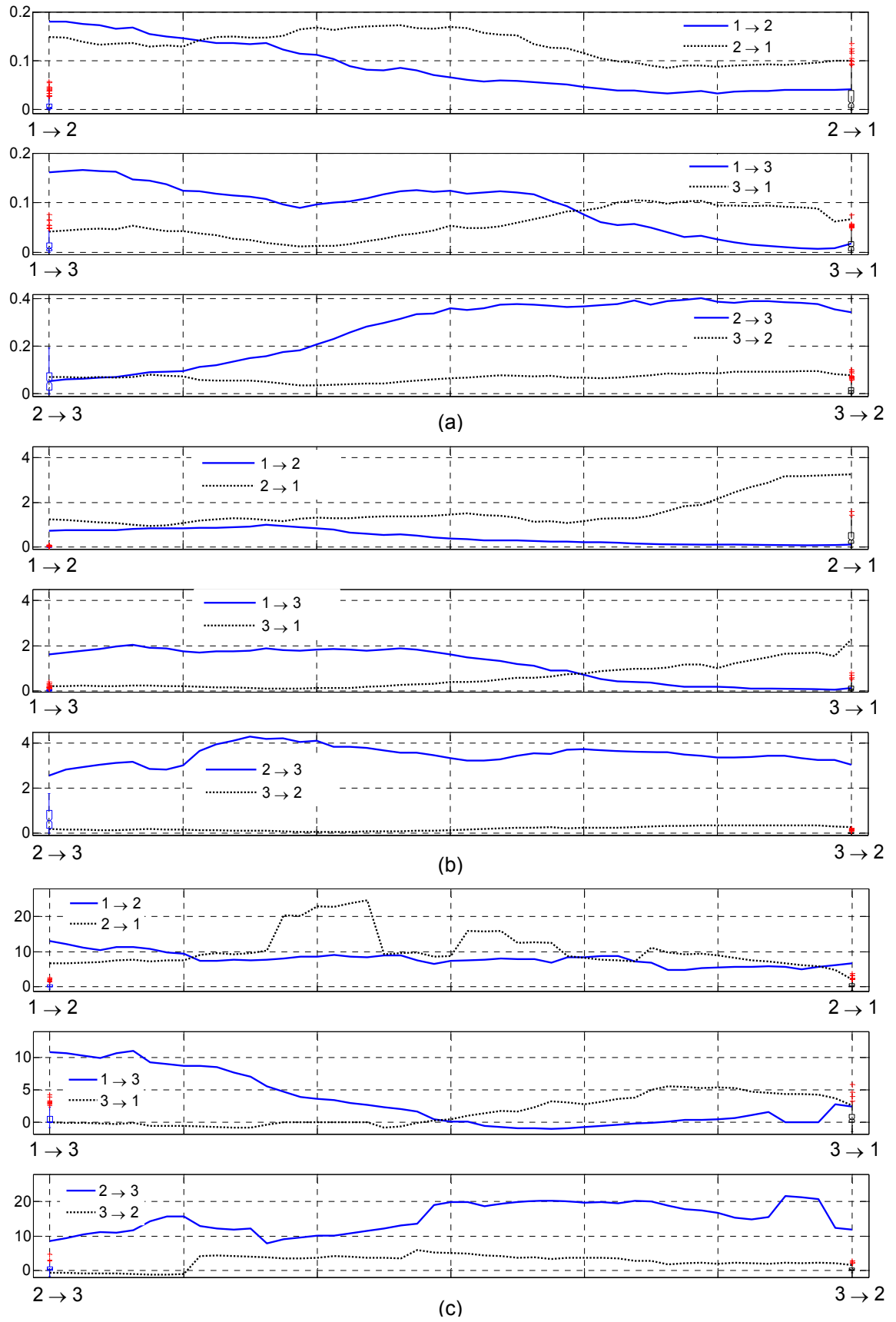
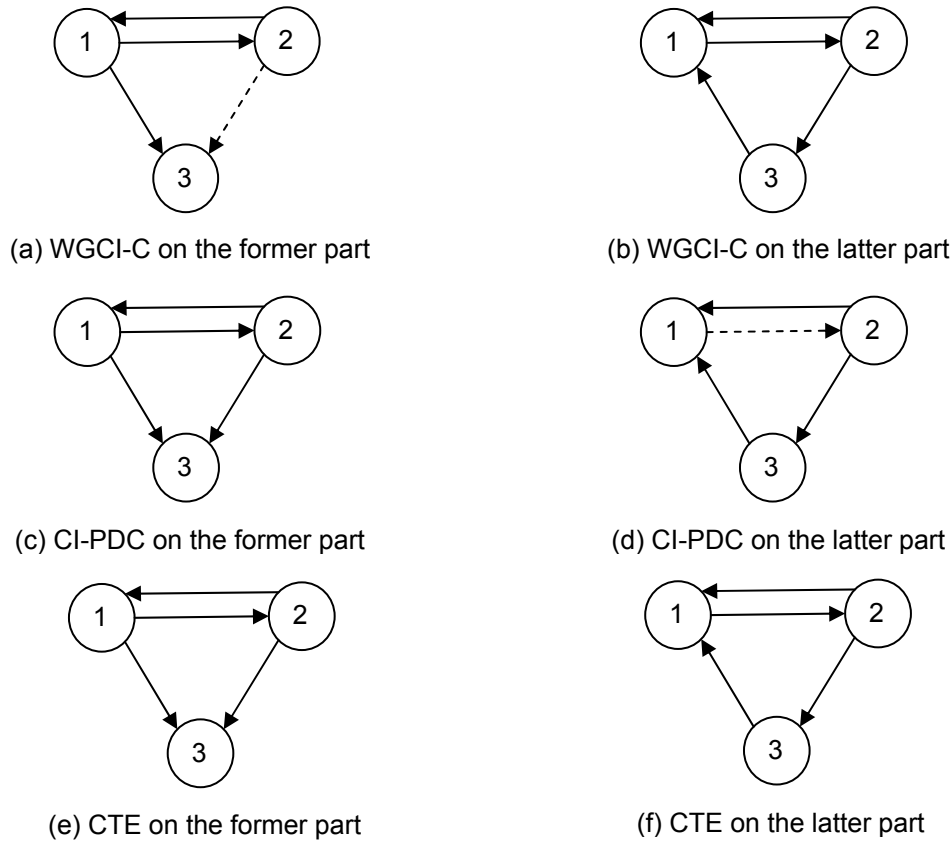


Fig. 5.33 – Results on WGCI-C, CI-PDC and CTE for 3 real signals.



*Fig. 5.34 – Estimated graphs from the results on WGCI-C, CI-PDC and CTE.
Solid line: relation of higher amplitude
Dashed line: relation of lower amplitude.*

5.5. Discussion and conclusion

In this chapter, we analyzed the performance of all the techniques described in chapter 4 to detect and identify various types of effective brain connectivity, *i.e.* unidirectional vs bidirectional and direct vs indirect causality relations, following the experimental protocol described in chapter 3. Firstly, these techniques are tested on Monte Carlo experiments using linear and nonlinear AR models. Performance is analyzed and compared in terms of mean values and standard deviation of the indices. First conclusions are summarized in Tab. 5.19. According to the experimental results, WGCI-C, CI-PDC and CTE are the most powerful techniques in their respective "family". Secondly, a validation of the results obtained by these three techniques has been conducted on simulated and real signals. For simulated AR models, these techniques proved excellent. Tested on a physiology-based model, WGCI-C and CTE seemed to be the most promising. A first application on signals recorded in a guinea-pig tends to reveal comparable trend.

Furthermore, CTE is a method which is very sensitive to the selection of the model order. The gBIC criterion improves the robustness of the CTE estimator. However, the optimization of the model order requires a longer computation time. In a future work, we plan to search how to save time.

According to the present results, if we adopt gBIC to estimate the order for WGCI-C, performance may be enhanced.

Note that, in the class of nonparametric and nonlinear methods, it should be possible to use nonlinear regression approach to test Granger causality. Such an approach should be an natural extension of the classical nonlinear and nonparametric correlation coefficient [Louis Dorr 2007, Wendling 2010]. Let us recall that this coefficient only searches for detecting a possible causality a functional relation between a past sample of the predictive channel and the present sample of the predicted channel, which clearly constitutes a coarse approach compared to that of Granger.

Method	unidirectional vs bidirectional relations	direct vs indirect relations
WGCI-P	Y	N
WGCI-C	Y	Y
FGCI-P	Y	N
FGCI-C	Y	Y
PSI-OC	N (just Y for unidirectional)	N
CI-PC	N (just Y for unidirectional)	Y (just for unidirectional)
CI-DCOH	Y	N
CI-DTF	Y	N
CI-PDC	Y	Y
TE	Y	N
CTE	Y	Y

Tab. 5.19 – Advantages and drawbacks of the techniques.

Y: discrimination between the types of relations under study.

N: no discrimination between the types of relations under study.

Conclusion

The objective of this work was to contribute to the analysis of brain connectivity and more specifically to that of effective connectivity, which, contrary to functional connectivity, allows for allocating a direction to a relation between two cerebral areas. If the interest in examining the relations between cerebral structures is no more to prove – "localisationist" and "globalist" visions, which conflicted until last century in associating any faculty to the activity of a particular structure or, on the opposite, to the activity of a network, agree today on a "connectivist" vision of the cerebral function (cognitive process associated to a global activity of a network of localized modules) [Mangin 1998] –, the methodology to reach this objective remains an open question. Even if this work may be viewed as a contribution in signal processing, effort has been done to describe the connectivity paradigm as well as the neurocomputational approach in neuroscience to understand the context of this study and its advances.

In this context, different approaches have been investigated to bring answers to two essential questions, *i.e.* the identification of unidirectional and/or bidirectional relations between structures and the discrimination between direct and indirect links. Our goal was to draw a state of the art of the main techniques found in the literature, propose some variants that could be useful and conduct an experimental study using Monte-Carlo experiments trying to derive tendencies. We constrained ourselves not to compare these techniques only on generic models. As a matter of fact, most of the techniques found in the literature deal with linear and nonlinear VAR models and also with nonlinear dynamic models generally relevant in chaotic systems (*e.g.* Lorenz and Rössler attractors). We chose to keep the first ones (linear and nonlinear VAR models) to get a kind of usual "core models" and to replace the second ones by physiology-based models (nonlinear models) to simulate intracerebral EEG signals.

The existing techniques devoted to connectivity analysis we selected first are those based on Granger causality index estimated in the time domain as well as its extended frequential and/or conditioned versions and also transfer entropy proposed as a nonlinear and nonparametric measure computed from a pair of signals. This method has also been considered in its conditional form to detect direct links taking into the presence of a third signal account. From the phase slope index introduced by Nolte, we proposed to develop it in the case of autoregressive modeling and to complete the search of directional links providing information on propagation delays between neuronal

ensembles. All techniques require the determination of N^2 prediction orders, if N is the number of channels considered simultaneously, either for the Granger approach relying on AR modeling or for entropic measures. The sensitivity of the measures to the order may be harmful, as seen experimentally with transfer entropy, an analysis of this sensitivity has been performed and a strategy to reduce computation time in the order estimation has been suggested.

Once calibrated, these methods revealed high-performance, it being understood that (i) only "conditional" or "partial" techniques (*i.e.* conditional to any other available information than that carried by the signals under study) are able to distinguish direct and indirect (*i.e.* mediated by other environmental signals) connectivity, and (ii) signals were generated without any additive noise, this hypothesis being widely accepted in the context of intracerebral EEG recordings. These techniques proved successful on simulations done on nonlinear AR signals and results corroborated those obtained in the linear case.

In a next step, we turned our study to physiology-based models, through the investigation of an existing macroscopic model of populations for which different sets of calibration parameters have to be searched for getting, on the one hand, comparable connectivity topologies as those described in the AR models, and, on the other hand, epileptic activities. In this step, only three techniques have been retained: the Granger causality index, the phase slope index based on partial directed coherence and the conditional transfer entropy. We first simulated a null hypothesis H_0 (absence of connectivity) using surrogate data (proposal of two methods) and we conducted a statistical analysis, using box plots and ROC curves, not only on the physiological model but also on the VAR models for a unique topology of connectivity (for time computing reasons). The indices remained robust to nonlinearities induced by these complex situations, and consistent graphs of propagation were retrieved confirming our first intuition made up by the previous results on nonlinear signals.

Finally, an analysis on real signals has been realized using signals recorded in a guinea-pig in three different structures (hippocampus, perirhinal cortex, medial entorhinal cortex) corresponding to a 10-s sequence. Without any ground-truth, which makes the assessment difficult, the use of surrogate data allowed us to speculate a good behavior of our techniques. Globally, for the three approaches, results appeared coherent.

What are the main conclusions of this work? Aware of challenges in brain connectivity, we had to answer the question of finding or adapting, on the basis of literature and experiments, a "core" of techniques giving the best compromise between performance and complexity. If our goal was extremely ambitious, we partially reached it, even if our contribution highlighted some weak points: if the nonlinear entropic approach seems attractive, its sensitivity to order estimation results in high computation time for a confident estimation. Moreover, the estimation of the probabilities distributions for high dimensions causes also trouble when memory allocation is not sufficient. If somebody is interested in measuring connectivity on much longer signals (to study predictability or long-term potentiation mechanisms, for example), we have to think about how to deal with this issue considering for example GPU (Graphics Processing Unit). In a more fundamental manner, we can "reconsider" the methods to estimate transfer entropy, or at least, verify whether some of them could improve the

results. Another difficulty – also linked to the order estimation – relies on the Markov hypothesis in Granger's approach. It is quite obvious that the outputs of our physiological models correspond to hidden Markov or ARMA processes and the basic concept of the causality measures estimated from these signals would need deeper investigation. The sensitivity of the techniques to additive disturbances in the signals' environment is also an open question.

Finally, concerning real signals, we must test a larger database (for instance on an animal model) to check whether applying our methods on this database is substantial and allows us to verify/suggest essential physiological hypotheses.

At this stage of our research, if some questions remain unanswered, we tried to contribute to knowledge on human cerebral connectivity. The interaction between different concepts that conflict sometimes and face each other – neuronal segregation and integration, influence and reciprocity, direct and mediated action – in neuronal cerebral networks produces diversified information, leading to complex models not completely controlled.

Beyond our methodological prospects to estimate information flow in propagation graphs, a better understanding of the organization of the epileptic networks could take advantage of multiple acquisition modalities optimizing temporal and spatial resolution, the finality being a finer delimitation of the cerebral volume to be resected to suppress epileptic activity (to cancel it or at least to stop it).

Appendix A

Power spectral densities of the signals

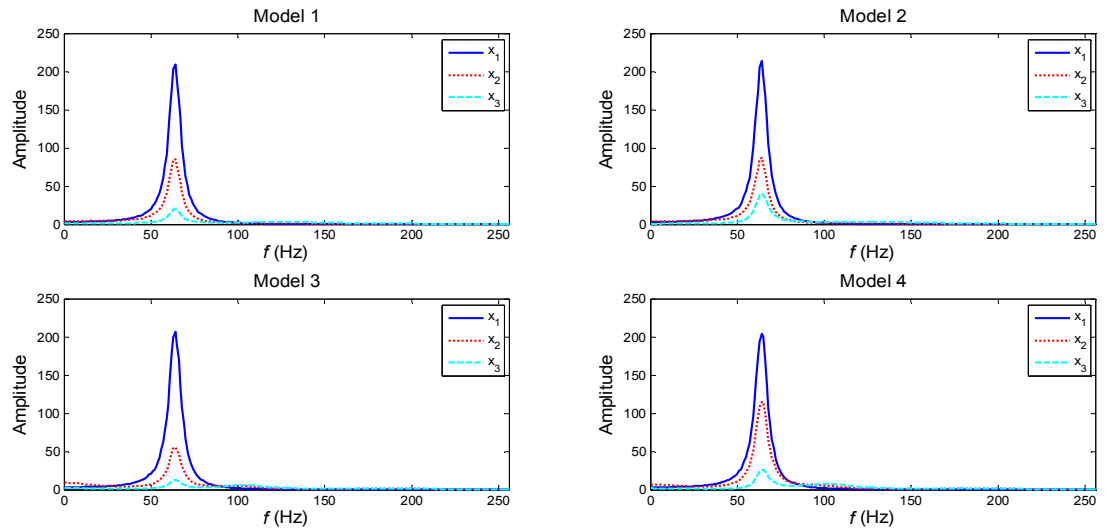


Fig. A.1 - PSD of the signals x_1 , x_2 and x_3 in the 4 linear AR models (see Eq. (3.1)).

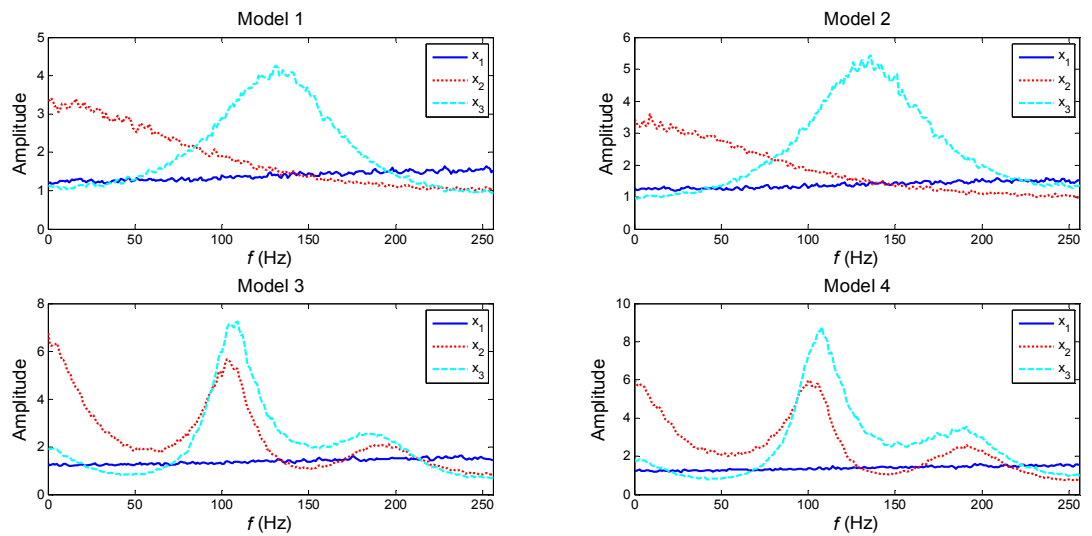


Fig. A.2 - PSD of the signals x_1 , x_2 and x_3 in the 4 nonlinear AR models (see Eq. (3.2)).

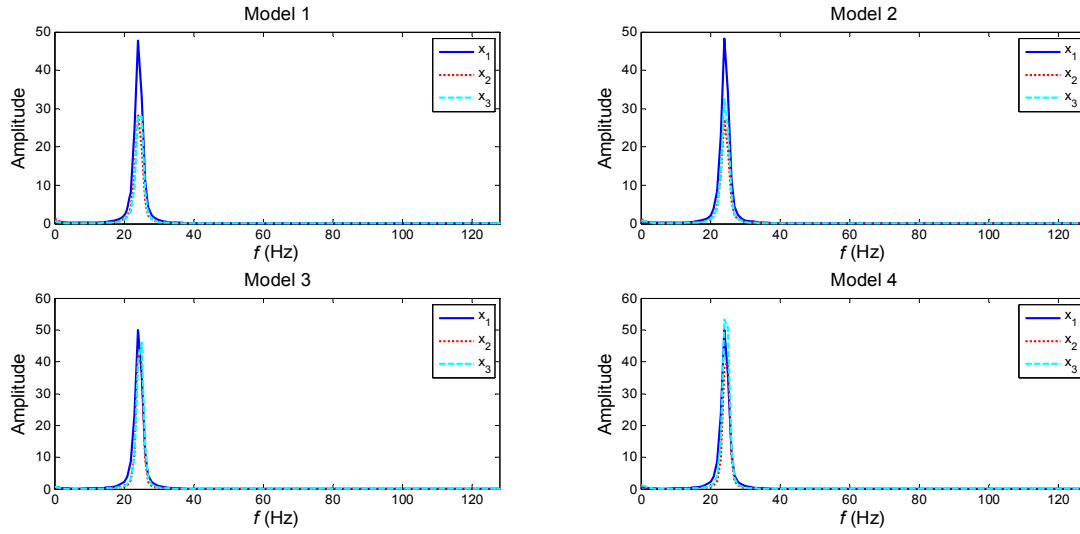


Fig. A.3 - PSD of the signals x_1 , x_2 and x_3 in the physiological models.

Model 1: $K^{12} = 1500$, $K^{23} = 1500$

Model 2: $K^{12} = 1500$, $K^{13} = 1500$ and $K^{23} = 1500$

Model 3: $K^{12} = 1500$, $K^{23} = 1500$ and $K^{32} = 1500$

Model 4: $K^{12} = 1500$, $K^{13} = 1500$, $K^{23} = 1500$ and $K^{32} = 1500$.

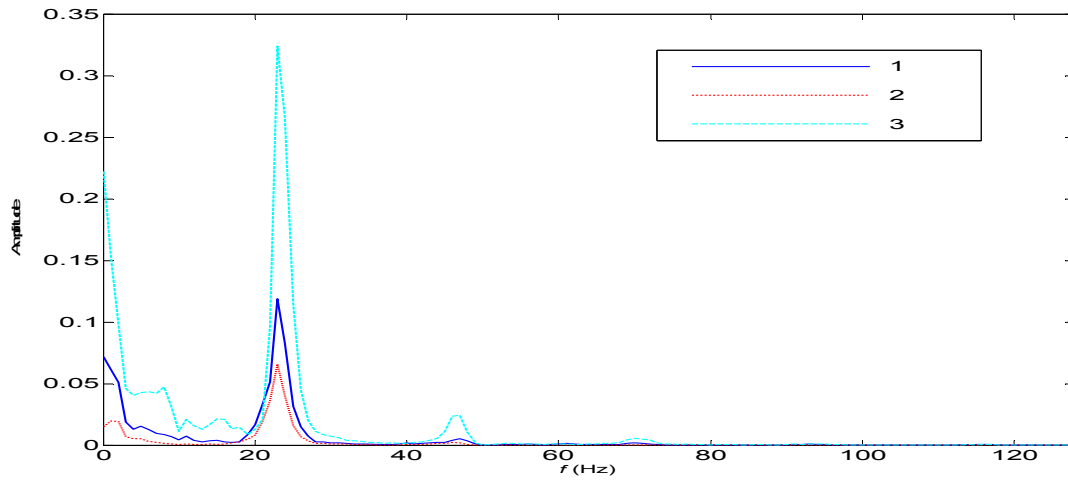


Fig. A.4 - PSD of the three recorded real signals.

Appendix B

Parameters of the physiological models

B.1. First scenario (model 1)

	Population 1				Population 2				Population 3			
	<i>A</i>	<i>B</i>	<i>G</i>	<i>g</i>	<i>A</i>	<i>B</i>	<i>G</i>	<i>g</i>	<i>A</i>	<i>B</i>	<i>G</i>	<i>g</i>
$K^{12} = 1500$ $K^{23} = 1500$	5	3	20	250	3.5	3.5	84	250	3.5	3.485	70.5	250
$K^{12} = 1000$ $K^{23} = 1000$	5	3	20	250	3.5	3.5	62	250	3.5	3.485	52.5	250
$K^{12} = 500$ $K^{23} = 500$	5	3	20	250	3.5	3.5	42.5	250	3.5	3.485	38	250

B.2. Second scenario (model 2)

	Population 1				Population 2				Population 3			
	<i>A</i>	<i>B</i>	<i>G</i>	<i>g</i>	<i>A</i>	<i>B</i>	<i>G</i>	<i>g</i>	<i>A</i>	<i>B</i>	<i>G</i>	<i>g</i>
$K^{12} = 1500$ $K^{13} = 1500$ $K^{23} = 1500$	5	3	20	250	3.5	3.5	84	250	3.5	1.3	65	250
$K^{12} = 1500$ $K^{13} = 750$ $K^{23} = 750$	5	3	20	250	3.5	3.5	84.4	250	3.5	1.3	40.5	250

B.3. Third scenario (model 3)

	Population 1				Population 2				Population 3			
	<i>A</i>	<i>B</i>	<i>G</i>	<i>g</i>	<i>A</i>	<i>B</i>	<i>G</i>	<i>g</i>	<i>A</i>	<i>B</i>	<i>G</i>	<i>g</i>
$K^{12} = 1500$	5	3	20	250	3.0	1	65	255	3.55	1	32.5	225
$K^{23} = 1500$												
$K^{32} = 1500$												

B.4. Fourth scenario (model 4)

	Population 1				Population 2				Population 3			
	<i>A</i>	<i>B</i>	<i>G</i>	<i>g</i>	<i>A</i>	<i>B</i>	<i>G</i>	<i>g</i>	<i>A</i>	<i>B</i>	<i>G</i>	<i>g</i>
$K^{12} = 1500$	5	3	20	250	3.0	1	66.5	225	3.55	1	57.3	225
$K^{13} = 1500$												
$K^{23} = 1500$												
$K^{32} = 1500$												

Appendix C

Typology of the methods

	Nonlinear nonparametric approach	Granger causality	AR modeling approach	Phase slope index	Linear nonparametric approach
Time	<div> <div> Ordi.: TE Cond.:CTE </div> <div> Ordi.: WGCi-P Cond.: WGCi-C </div> </div>				
Frequency	<div> <div> Ordi.: FGCI-P (Geweke) Cond.: FGCI-C (Geweke) </div> <div> Ordi.: PSI-OC Cond.: CI-PC </div> <div> Ordi.: CI-DCOH (2 channels) CI-DTF (Q channels) Cond.: CI-PDC </div> </div>				

This table groups all the proposed indices together according to their representation domain (time or frequency) and to the main concepts they are based on. (Ordi.: Ordinary, Cond. Conditional). The indices in blue are those selected for the statistical analysis.

Appendix D

Determining the model order (AIC, BIC)

D.1. Introduction

The Akaike Information Criterion (AIC) proposed by Akaike in 1973 [Akaike 1973] is a classical linear measure to choose the autoregressive (AR) model order. As an alternative criterion, the Bayesian Information Criterion (BIC) was proposed by Schwarz in 1978 [Schwarz 1978]. Except the application of selecting the AR model order, AIC and BIC have also been applied to estimate the order of Markov chain [Katz 1981, Tong 1975]. In the following, the classical AIC and BIC procedures are firstly introduced with the application of selecting the AR model order, then the modified AIC and BIC definitions are presented with the goal of obtaining the optimal order of AR model and Markov process.

D.2. Akaike's information criterion and the bayesian information criterion

Given a real m -dimensional vectorial AR (VAR) process $z_n = [z_n(1), z_n(2), \dots, z_n(m)]^T$ with zero mean

$$z_n = \Phi_1 z_{n-1} + \Phi_2 z_{n-2} + \dots + \Phi_p z_{n-p} + w_n \quad (D.1)$$

where $\Phi_i, i = 1, 2, \dots, p$ are $m \times m$ coefficient matrices and p the model order. The process w_n is a zero mean independent identically distributed vector process with covariance matrix Σ . It is assumed that w_n and z_{n-i} are independent for each variable $i > 0$, and that z_n is ergodic in the wide sense (ergodic in the first and second moments).

Suppose that the observations z_1, z_2, \dots, z_N are generated by the VAR process given in Eq. (D.1). The least-squares method [Brockwell 1991] can be used to fit a VAR model with order q to the observations. Suppose that the estimated model is as follows:

$$z_n = \hat{\Phi}_1 z_{n-1} + \hat{\Phi}_2 z_{n-2} + \dots + \hat{\Phi}_q z_{n-q} + \hat{w}_n. \quad (D.2)$$

The coefficient matrices $\hat{\Phi}_i, i = 1, 2, \dots, q$ of this estimated model are obtained by solving the following set of equations:

$$ZA_j = z_{q+1,j} \quad (j = 1, 2, \dots, m) \quad (D.3)$$

where $Z = \begin{bmatrix} z_q^T & z_{q-1}^T & \cdots & z_1^T \\ z_{q+1}^T & z_q^T & \cdots & z_2^T \\ \vdots & \vdots & \ddots & \vdots \\ z_{N-1}^T & z_{N-2}^T & \cdots & z_{N-q}^T \end{bmatrix}$, $z_{q+1,j} = \begin{bmatrix} z_{q+1}(j) \\ z_{q+2}(j) \\ \vdots \\ z_N(j) \end{bmatrix}$, and $A_j = \begin{bmatrix} \Phi_1^T(j,:) \\ \Phi_2^T(j,:) \\ \vdots \\ \Phi_q^T(j,:) \end{bmatrix}$, $\Phi_i(j,:)$ is the j^{th} row

of Φ_i . The least-squares solution of Eq. (D.3) is given by

$$\hat{A}_j = (Z^T Z)^{-1} Z^T z_{q+1,j} \quad (j = 1, 2, \dots, m). \quad (D.4)$$

Then, the residual covariance matrix of the input noise of the VAR model can be estimated as

$$\hat{\Sigma} = \frac{1}{N-q} \sum_{n=q+1}^N (z_n - \hat{z}_n)(z_n - \hat{z}_n)^T \quad (D.5)$$

where $\hat{z}_n = \hat{\Phi}_1 z_{n-1} + \hat{\Phi}_2 z_{n-2} + \cdots + \hat{\Phi}_q z_{n-q}$.

The above estimation procedure can be carried out for any model order q . The optimal order q is usually determined by minimizing the AIC defined by:

$$AIC(q) = N \ln(\det(\hat{\Sigma})) + 2m^2 q. \quad (D.6)$$

Plotted as a function of q , the proper model order corresponds to the minimum of this function.

An alternative criterion is the BIC, which is defined by:

$$BIC(q) = N \ln(\det(\hat{\Sigma})) + m^2 q \ln(N). \quad (D.7)$$

Appendix E

Notations

Method	Chapter 3	Chapter 4	Chapter 5
WGCI-P	$Ind_{X_i \rightarrow X_j}$	$WGCI_{X_i \rightarrow X_j}$ -P (Eqs. (4.5) and (4.6))	$WGCI_{i \rightarrow j}$ -P
WGCI-C	$Ind_{X_i \rightarrow X_j / X_k}$	$WGCI_{X_i \rightarrow X_j X_k}$ -C (Eq. (4.8))	$WGCI_{i \rightarrow j}$ -C
FGCI-P	$Ind_{X_i \rightarrow X_j}$	$FGCI_{X_i \rightarrow X_j}(f)$ -P (Eqs. (4.18) and (4.21))	$FGCI_{i \rightarrow j}$ -P
FGCI-C	$Ind_{X_i \rightarrow X_j / X_k}$	$FGCI_{X_i \rightarrow X_j X_k}(f)$ -C (Eq. (4.37))	$FGCI_{i \rightarrow j}$ -C
PSI-OC	$Ind_{X_i \rightarrow X_j}$	PSI_{ij} -OC (Eq. (4.52))	PSI_{ij} -OC
CI-PC	$Ind_{X_i \rightarrow X_j / X_k}$	CI_{ij} -PC (Eq. (4.62))	CI_{ij} -PC
CI-DCOH	$Ind_{X_i \rightarrow X_j}$	CI_{ij} -DCOH (Eq. (4.67))	CI_{ij} -DCOH
CI-DTF	$Ind_{X_i \rightarrow X_j}$	CI_{ij} -DTF (Eq. (4.69))	CI_{ij} -DTF
CI-PDC	$Ind_{X_i \rightarrow X_j / X_k}$	CI_{ij} -PDC (Eq. (4.71))	CI_{ij} -PDC
TE	$Ind_{X_i \rightarrow X_j}$	$TE_{Y \rightarrow X}$ (Eq. (4.78))	$TE_{i \rightarrow j}$
CTE	$Ind_{X_i \rightarrow X_j / X_k}$	$CTE_{Y \rightarrow X Z}$ (Eq. (4.88))	$CTE_{i \rightarrow j k}$

References

- [Achard 2006] Achard, S., Salvador, R., Whitcher, B., Suckling, J., and Bullmore, E., "A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs," *The Journal of Neuroscience*, vol. 26, no. 1, pp. 63-72, 2006.
- [Achard 2007] Achard, S., and Bullmore, E., "Efficiency and cost of economical brain functional networks," *PLoS Computational Biology*, vol. 3, no. 2, pp. (e17) 174-183, 2007.
- [Achard 2008] Achard, S., Bassett, D. S., Meyer-Lindenberg, A., and Bullmore, E., "Fractal connectivity of long-memory networks," *Physical Review E*, vol. 77, no. 3, pp. 036104-1:036104-12, 2008.
- [Adcock 2003] Adcock, J. E., Wise, R. G., Oxbury, J. M., Oxbury, S. M., and Matthews, P. M., "Quantitative fMRI assessment of the differences in lateralization of language-related brain activation in patients with temporal lobe epilepsy," *NeuroImage*, vol. 18, no. 2, pp. 423-438, 2003.
- [Akaike 1973] Akaike, H., "Information theory and an extension of the maximum likelihood principle," *the Second International Symposium on Information Theory, Budapest, Hungary*. pp. 267-281, 2-8 September 1973.
- [Ansari-Asl 2006] Ansari-Asl, K., Senhadji, L., Bellanger, J. J., and Wendling, F., "Quantitative evaluation of linear and nonlinear methods characterizing interdependencies between brain signals," *Physical Review E*, vol. 74, no. 3, pp. 031916, 2006.
- [Ansari-Asl 2007] Ansari-Asl, K., Chanel, G., and Pun, T., "A channel selection method for EEG classification in emotion assessment based on synchronization likelihood," 2007.
- [Arnhold 1999] Arnhold, J., Grassberger, P., Lehnertz, K., and Elger, C. E., "A robust method for detecting interdependences: application to intracranially recorded EEG," *Physica D: Nonlinear Phenomena*, vol. 134, no. 4, pp. 419-430, 1999.
- [Astolfi 2006] Astolfi, L., Cincotti, F., Mattia, D., Marciani, M. G., Baccalá, L. A., De Vico, F. F., Salinari, S., Ursino, M., Zavaglia, M., and Babiloni, F., "Assessing cortical functional connectivity by partial directed coherence: simulations and application to real data," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 9, pp. 1802-1812, 2006.
- [Babb 1989] Babb, T. L., Pretorius, J. K., Kupfer, W. R., and Crandall, P. H., "Glutamate decarboxylase-immunoreactive neurons are preserved in human epileptic hippocampus," *The Journal of Neuroscience*, vol. 9, no. 7, pp. 2562-2574, 1989.
- [Baccalá 2001] Baccalá, L. A., and Sameshima, K., "Partial directed coherence: a new concept in neural structure determination," *Biological Cybernetics*, vol. 84, no. 6, pp. 463-474, 2001.

- [Baker 1997] Baker, G. A., Jacoby, A., Buck, D., Stalgis, C., and Monnet, D., "Quality of life of people with epilepsy: a European study," *Epilepsia*, vol. 38, no. 3, pp. 353-362, 1997.
- [Baker 2000] Baker, G. A., Brooks, J., Buck, D., and Jacoby, A., "The stigma of epilepsy: a European perspective," *Epilepsia*, vol. 41, no. 1, pp. 98-104, 2000.
- [Bal 2000] Bal, T., Debay, D., and Destexhe, A., "Cortical feedback controls the frequency and synchrony of oscillations in the visual thalamus," *The Journal of Neuroscience*, vol. 20, no. 19, pp. 7478-7488, 2000.
- [Banerjee 2009] Banerjee, P. N., Filippi, D., and Allen Hauser, W., "The descriptive epidemiology of epilepsy--A review," *Epilepsy Research*, vol. 85, no. 1, pp. 31-45, 2009.
- [Barnett 2009] Barnett, L., Barrett, A. B., and Seth, A. K., "Granger causality and transfer entropy are equivalent for Gaussian variables," *Physical Review Letters*, vol. 103, no. 23, pp. 238701, 2009.
- [Basser 1994] Basser, P. J., Mattiello, J., and Lebihan, D., "MR diffusion tensor spectroscopy and imaging," *Biophysical Journal*, vol. 66, no. 1, pp. 259-267, 1994.
- [Bassett 2006a] Bassett, D. S., and Bullmore, E., "Small-world brain networks," *The Neuroscientist*, vol. 12, no. 6, pp. 512-523, 2006a.
- [Bassett 2006b] Bassett, D. S., Meyer-Lindenberg, A., Achard, S., Duke, T., and Bullmore, E., "Adaptive reconfiguration of fractal small-world human brain functional networks," *Proceedings of the National Academy of Sciences*, vol. 103, no. 51, pp. 19518-19523, 2006b.
- [Bendat 1986a] Bendat, J. S., and Piersol, A. G., *Random Data. Analysis and Measurement Procedures*, 2nd ed., New York: John Wiley & Sons, 1986a.
- [Bendat 1986b] Bendat, J. S., and Piersol, A. G., "Decomposition of wave forces into linear and non-linear components," *Journal of Sound Vibration*, vol. 106, pp. 391-408, 1986b.
- [Benke 2006] Benke, T., Köylü, B., Visani, P., Karner, E., Brenneis, C., Bartha, L., Trinkä, E., Trieb, T., Felber, S., and Bauer, G., "Language lateralization in temporal lobe epilepsy: a comparison between fMRI and the Wada Test," *Epilepsia*, vol. 47, no. 8, pp. 1308-1319, 2006.
- [Berg 2007] Berg, A. T., Langfitt, J. T., Spencer, S. S., and Vickrey, B. G., "Stopping antiepileptic drugs after epilepsy surgery: a survey of US epilepsy center neurologists," *Epilepsy & Behavior*, vol. 10, no. 2, pp. 219-222, 2007.
- [Besserve 2011] Besserve, M., and Martinerie, J., "Extraction of functional information from ongoing brain electrical activity," *IRBM*, vol. 32, no. 1, pp. 27-34, 2011.
- [Bestmann 2008] Bestmann, S., Ruff, C. C., Blankenburg, F., Weiskopf, N., Driver, J., and Rothwell, J. C., "Mapping causal interregional influences with concurrent TMS-fMRI," *Experimental Brain Research*, vol. 191, no. 4, pp. 383-402, 2008.
- [Bhattacharya 2001] Bhattacharya, J., "Reduced degree of long-range phase synchrony in pathological human brain," *Acta neurobiologiae experimentalis*, vol. 61, no. 4, pp. 309-318, 2001.
- [Billingsley 1965] Billingsley, P., *Ergodic Theory and Information*, New York: Wiley, 1965.

- [Birbeck 2002] Birbeck, G. L., Hays, R. D., Cui, X., and Vickrey, B. G., "Seizure reduction and quality of life improvements in people with epilepsy," *Epilepsia*, vol. 43, no. 5, pp. 535-538, 2002.
- [Blalock 1961] Blalock, H. M., *Causal Inferences in Nonexperimental Research*, Chapel Hill, NC: University of North Carolina Press, 1961.
- [Blinowska 2004] Blinowska, K. J., Kuś, R., and Kamiński, M., "Granger causality and information flow in multivariate processes," *Physical Review E*, vol. 70, no. 5, pp. 050902, 2004.
- [Blume 2001] Blume, W. T., Lüders, H. O., Mizrahi, E., Tassinari, C., Van Emde Boas, W., and Engel Jr, E., "Glossary of descriptive terminology for ictal semiology: report of the ILAE task force on classification and terminology," *Epilepsia*, vol. 42, no. 9, pp. 1212-1218, 2001.
- [Boccaletti 2006] Boccaletti, S., Hwang, D. U., Chávez, M., Amann, A., Kurths, J., and Pecora, L. M., "Synchronization in dynamical networks: Evolution along commutative graphs," *Physical Review E*, vol. 74, no. 1, pp. 016102-1:016102-5, 2006.
- [Boutillier 2007a] Boutillier, B., and Outrequin, G., 2007a, <http://www.anatomie-humaine.com/Le-Cerveau-1.html>.
- [Boutillier 2007b] Boutillier, B., and Outrequin, G., 2007b, <http://www.anatomie-humaine.com/Presentation-Generalites.html>.
- [Brandes 2005] Brandes, U., and Erlebach, T., *Network Analysis: Methodological Foundations*: Springer Verlag, 2005.
- [Brazier 1952] Brazier, M. A. B., and Casby, J. U., "Crosscorrelation and autocorrelation studies of electroencephalographic potentials," *Electroencephalography and Clinical Neurophysiology*, vol. 4, no. 2, pp. 201-211, 1952.
- [Brazier 1968] Brazier, M. A. B., "Studies of the EEG activity of limbic structures in man," *Electroencephalography and Clinical Neurophysiology*, vol. 25, no. 4, pp. 309-318, 1968.
- [Breakspear 2004] Breakspear, M., Brammer, M. J., Bullmore, E. T., Das, P., and Williams, L. M., "Spatiotemporal wavelet resampling for functional neuroimaging data," *Human Brain Mapping*, vol. 23, no. 1, pp. 1-25, 2004.
- [Bressler 2007] Bressler, S. L., Richter, C. G., Chen, Y., and Ding, M., "Cortical functional network organization from autoregressive modeling of local field potential oscillations," *Statistics in Medicine*, vol. 26, no. 21, pp. 3875-3885, 2007.
- [Bressler 2011] Bressler, S. L., and Seth, A. K., "Wiener-Granger Causality: A well established methodology," *NeuroImage*, vol. 58, no. 2, pp. 323-329, 2011.
- [Brockwell 1991] Brockwell, P. J., and Davis, R. A., *Time Series: Theory and Methods*, second ed.: Springer Verlag, 1991.
- [Brodie 2005] Brodie, M. J., "Diagnosing and predicting refractory epilepsy," *Acta Neurologica Scandinavica*, vol. 112, pp. 36-39, 2005.
- [Brovelli 2004] Brovelli, A., Ding, M., Ledberg, A., Chen, Y., Nakamura, R., and Bressler, S. L., "Beta oscillations in a large-scale sensorimotor cortical network: directional influences revealed by Granger causality," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 26, pp. 9849-9854, 2004.

- [Cao 1997] Cao, L., "Practical method for determining the minimum embedding dimension of a scalar time series," *Physica D: Nonlinear Phenomena*, vol. 110, no. 1-2, pp. 43-50, 1997.
- [Cascino 1994] Cascino, G. D., "Epilepsy: contemporary perspectives on evaluation and treatment," *Mayo Clinic Proceedings, Rochester, Minn.* pp. 1199-1211, 1994.
- [Chávez 2003] Chávez, M., Martinerie, J., and Le Van Quyen, M., "Statistical assessment of nonlinear causality: application to epileptic EEG signals," *Journal of Neuroscience Methods*, vol. 124, no. 2, pp. 113-128, 2003.
- [Chávez 2005] Chávez, M., Hwang, D. U., Amann, A., Hentschel, H. G. E., and Boccaletti, S., "Synchronization is Enhanced in Weighted Complex Networks," *Physical Review Letters*, vol. 94, no. 21, pp. 218701-1:218701-4, 2005.
- [Chávez 2006a] Chávez, M., Besserve, M., Adam, C., and Martinerie, J., "Towards a proper estimation of phase synchronization from time series," *Journal of Neuroscience Methods*, vol. 154, no. 1-2, pp. 149-160, 2006a.
- [Chávez 2006b] Chávez, M., Hwang, D. U., Amann, A., and Boccaletti, S., "Synchronizing weighted complex networks," *Chaos (Woodbury, NY)*, vol. 16, no. 1, pp. 015106-1:015106-7, 2006b.
- [Chávez 2007] Chávez, M., Hwang, D. U., and Boccaletti, S., "Synchronization processes in complex networks," *The European Physical Journal-Special Topics*, vol. 146, no. 1, pp. 129-144, 2007.
- [Chen 2004] Chen, Y., Rangarajan, G., Feng, J., and Ding, M., "Analyzing multiple nonlinear time series with extended Granger causality," *Physics Letters A*, vol. 324, no. 1, pp. 26-35, 2004.
- [Chen 2006] Chen, Y., Bressler, S. L., and Ding, M., "Frequency decomposition of conditional Granger causality and application to multivariate neural field potential data," *Journal of Neuroscience Methods*, vol. 150, no. 2, pp. 228-237, 2006.
- [Chicharro 2012] Chicharro, D., "On the spectral formulation of Granger causality," *Biological Cybernetics*, vol. 105, no. 5-6, pp. 331-347, 2012.
- [Cooray 2011] Cooray, G. K., Hyllienmark, L., and Brismar, T., "Decreased cortical connectivity and information flow in type 1 diabetes," *Clinical Neurophysiology*, vol. 122, no. 10, pp. 1943-1950, 2011.
- [Cover 1991] Cover, T. M., and Thomas, J. A., *Elements of Information Theory*, New York: John Wiley & Sons, 1991.
- [Daly 1968] Daly, D. D., "Reflections on the concept of petit mal," *Epilepsia*, vol. 9, no. 3, pp. 175-178, 1968.
- [Daunizeau 2009] Daunizeau, J., Kiebel, S. J., and Friston, K. J., "Dynamic causal modelling of distributed electromagnetic responses," *NeuroImage*, vol. 47, no. 2, pp. 590-601, 2009.
- [David 2008] David, O., Guillemain, I., Sallet, S., Reyt, S., Deransart, C., Segebarth, C., and Depaulis, A., "Identifying neural drivers with functional MRI: an electrophysiological validation," *PLoS Biology*, vol. 6, no. 12, pp. 2683-2697, 2008.
- [Delorme 2002] Delorme, A., Makeig, S., Fabre-Thorpe, M., and Sejnowski, T., "From single-trial EEG to brain area dynamics," *Neurocomputing*, vol. 44-46, pp. 1057-1064, 2002.

- [Destexhe 1996] Destexhe, A., Bal, T., McCormick, D. A., and Sejnowski, T. J., "Ionic mechanisms underlying synchronized oscillations and propagating waves in a model of ferret thalamic slices," *Journal of Neurophysiology*, vol. 76, no. 3, pp. 2049-2070, 1996.
- [Dhamala 2008] Dhamala, M., Rangarajan, G., and Ding, M., "Analyzing information flow in brain networks with nonparametric Granger causality," *NeuroImage*, vol. 41, no. 2, pp. 354-362, 2008.
- [Ding 2006] Ding, M., Chen, Y., and Bressler, S. L., *Granger causality: basic theory and application to neuroscience*, Berlin: Wiley-VCH, 2006.
- [Duncan 2006] Duncan, J. S., Sander, J. W., Sisodiya, S. M., and Walker, M. C., "Adult epilepsy," *The Lancet*, vol. 367, no. 9516, pp. 1087-1100, 2006.
- [E.F.A. 2012] E.F.A. "Epilepsy Syndromes," Epilepsy Foundation of America; <http://www.epilepsyfoundation.org/about/types/syndromes/index.cfm>.
- [E.P.F.L.] E.P.F.L., <http://bluebrain.epfl.ch/>.
- [Edeline 1999] Edeline, J. M., "Learning-induced physiological plasticity in the thalamo-cortical sensory systems: a critical evaluation of receptive field plasticity, map changes and their potential mechanisms," *Progress in Neurobiology*, vol. 57, no. 2, pp. 165-224, 1999.
- [Eichler 2006] Eichler, M., "On the evaluation of information flow in multivariate systems by the directed transfer function," *Biological Cybernetics*, vol. 94, no. 6, pp. 469-482, 2006.
- [Engel 1996] Engel, J., "Surgery for seizures," *The New England Journal of Medicine*, vol. 334, no. 10, pp. 647-652, 1996.
- [Engel 2001] Engel, J., "Classification of epileptic disorders," *Epilepsia*, vol. 42, no. 3, pp. 316-316, 2001.
- [Engel 2006] Engel, J., "ILAE classification of epilepsy syndromes," *Epilepsy Research*, vol. 70, pp. 5-10, 2006.
- [Ermentrout 1998] Ermentrout, B., "Neural networks as spatio-temporal pattern-forming systems," *Reports on Progress in Physics*, vol. 61, no. 4, pp. 353-430, 1998.
- [Everitt 1998] Everitt, A. D., and Sander, J. W., "Incidence of epilepsy is now higher in elderly people than children," *BMJ: British Medical Journal*, vol. 316, no. 7133, pp. 780, 1998.
- [Fan 1995] Fan, J., and Gijbels, I., "Data-driven bandwidth selection in local polynomial fitting: variable bandwidth and spatial adaptation," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 57, no. 2, pp. 371-394, 1995.
- [Fauci 2008] Fauci, A. S., Braunwald, E., Kasper, D. L., Hauser, S. L., Longo, D. L., Jameson, J. L., and Loscalzo, J., *Harrison's Principles of Internal Medicine*: McGraw-Hill Professional, 2008.
- [Faugeras 2009] Faugeras, O., Touboul, J., and Cessac, B., "A constructive mean-field analysis of multi-population neural networks with random synaptic weights and stochastic inputs," *Frontiers in Computational Neuroscience*, vol. 3, no. 1, pp. 1-28, 2009.
- [Ferrerri 2011] Ferreri, F., Pasqualetti, P., Määttä, S., Ponzo, D., Ferrarelli, F., Tononi, G., Mervaala, E., Miniussi, C., and Rossini, P. M., "Human brain connectivity during

single and paired pulse transcranial magnetic stimulation," *NeuroImage*, vol. 54, no. 1, pp. 90-102, 2011.

- [Fingelkurts 2005] Fingelkurts, A. A., Fingelkurts, A. A., and Kähkönen, S., "Functional connectivity in the brain - is it an elusive concept?," *Neuroscience & Biobehavioral Reviews*, vol. 28, no. 8, pp. 827-836, 2005.
- [Fisher 2005] Fisher, R. S., Boas, W. E., Blume, W., Elger, C., Genton, P., Lee, P., and Engel Jr, J., "Epileptic seizures and epilepsy: definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE)," *Epilepsia*, vol. 46, no. 4, pp. 470-472, 2005.
- [Flechsig 1920] Flechsig, P. E., *Anatomie des Menschlichen Gehirns und Rückenmarks auf Myelogenetischer Grundlage*: G. Thieme, 1920.
- [Forsgren 2005] Forsgren, L., Beghi, E., Öun, A., and Sillanpää, M., "The epidemiology of epilepsy in Europe - a systematic review," *European Journal of Neurology*, vol. 12, no. 4, pp. 245-253, 2005.
- [Franaszczuk 1998] Franaszczuk, P. J., and Bergey, G. K., "Application of the directed transfer function method to mesial and lateral onset temporal lobe seizures," *Brain topography*, vol. 11, no. 1, pp. 13-21, 1998.
- [Fraser 1986] Fraser, A. M., and Swinney, H. L., "Independent coordinates for strange attractors from mutual information," *Physical Review A*, vol. 33, no. 2, pp. 1134, 1986.
- [Freeman 1973] Freeman, W. J., "A model of the olfactory system," *Neural Modeling*, vol. 1, pp. 41-62, 1973.
- [Freeman 1975] Freeman, W. J., *Mass Action in the Nervous System*, New York: Academic Press, 1975.
- [Freeman 1987] Freeman, W. J., "Simulation of chaotic EEG patterns with a dynamic model of the olfactory system," *Biological Cybernetics*, vol. 56, no. 2, pp. 139-150, 1987.
- [French 2007] French, J. A., "Refractory epilepsy: clinical overview," *Epilepsia*, vol. 48, pp. 3-7, 2007.
- [Frenzel 2007] Frenzel, S., and Pompe, B., "Partial mutual information for coupling analysis of multivariate time series," *Physical Review Letters*, vol. 99, no. 20, pp. 204101, 2007.
- [Friston 1994] Friston, K. J., "Functional and effective connectivity in neuroimaging: a synthesis," *Human Brain Mapping*, vol. 2, no. 1-2, pp. 56-78, 1994.
- [Friston 2003] Friston, K. J., Harrison, L., and Penny, W., "Dynamic causal modelling," *NeuroImage*, vol. 19, no. 4, pp. 1273-1302, 2003.
- [Friston 2009a] Friston, K. J., "Causal modelling and brain connectivity in functional magnetic resonance imaging," *PLoS Biology*, vol. 7, no. 2, pp. e1000033, 2009a.
- [Friston 2009b] Friston, K. J., "Modalities, modes, and models in functional neuroimaging," *Science*, vol. 326, no. 5951, pp. 399-403, 2009b.
- [Friston 2010] Friston, K. J., and Dolan, R. J., "Computational and dynamic models in neuroimaging," *NeuroImage*, vol. 52, no. 3, pp. 752-765, 2010.

- [Frogerais 2008] Frogerais, P., "Modélisation et identification en épilepsie : De la dynamique des populations neuronales aux signaux EEG," LTSI - Laboratoire Traitement du Signal et de l'Image, Université Rennes 1, Rennes, France, 2008.
- [Gaitatzis 2004] Gaitatzis, A., and Sander, J. W., "The mortality of epilepsy revisited," *Epileptic Disorders*, vol. 6, pp. 3-14, 2004.
- [Geweke 1982] Geweke, J. F., "Measurement of linear dependence and feedback between multiple time series," *Journal of the American Statistical Association*, vol. 77, no. 378, pp. 304-313, 1982.
- [Geweke 1984] Geweke, J. F., "Measures of conditional linear dependence and feedback between time series," *Journal of the American Statistical Association*, vol. 79, no. 388, pp. 907-915, 1984.
- [Gourévitch 2006] Gourévitch, B., Le-Bouquin-Jeannès, R., and Faucon, G., "Linear and nonlinear causality between signals: methods, examples and neurophysiological applications," *Biological Cybernetics*, vol. 95, no. 4, pp. 349-369, 2006.
- [Granger 1969] Granger, C. W. J., "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*, vol. 37, no. 3, pp. 424-438, Aug, 1969.
- [Grassberger 1983] Grassberger, P., and Procaccia, I., "Measuring the strangeness of strange attractors," *Physica D: Nonlinear Phenomena*, vol. 9, no. 1-2, pp. 189-208, 1983.
- [Grech 2008] Grech, R., Cassar, T., Muscat, J., Camilleri, K., Fabri, S., Zervakis, M., Xanthopoulos, P., Sakkalis, V., and Vanrumste, B., "Review on solving the inverse problem in EEG source analysis," *Journal of Neuroengineering and Rehabilitation*, vol. 5, no. 1, pp. 25: 1-33, 2008.
- [Guo 2008] Guo, S., Seth, A. K., Kendrick, K. M., Zhou, C., and Feng, J., "Partial Granger causality - Eliminating exogenous inputs and latent variables," *Journal of Neuroscience Methods*, vol. 172, no. 1, pp. 79-93, 2008.
- [Hagmann 2007] Hagmann, P., Kurrant, M., Gigandet, X., Thiran, P., Wedeen, V. J., Meuli, R., and Thiran, J. P., "Mapping human whole-brain structural networks with diffusion MRI," *PLoS One*, vol. 2, no. 7, pp. e597, 2007.
- [Hallett 2000] Hallett, M., "Transcranial magnetic stimulation and the human brain," *Nature*, vol. 406, no. 6792, pp. 147-150, 2000.
- [Haufe 2010] Haufe, S., Tomioka, R., Nolte, G., Müller, K. R., and Kawanabe, M., "Modeling sparse connectivity between underlying brain sources for EEG/MEG," *IEEE Transactions on Biomedical Engineering* vol. 57, no. 8, pp. 1954-1963, 2010.
- [Haufe 2011] Haufe, S., Nikulin, V., and Nolte, G., "Identifying brain effective connectivity patterns from EEG: performance of Granger Causality, DTF, PDC and PSI on simulated data," *BMC Neuroscience*, vol. 12, no. Suppl 1, pp. P141, 2011.
- [He 1998] He, J., and Hashikawa, T., "Connections of the dorsal zone of cat auditory cortex," *Journal of Comparative Neurology*, vol. 400, no. 3, pp. 334-348, 1998.
- [Hesse 2003] Hesse, W., Möller, E., Arnold, M., and Schack, B., "The use of time-variant EEG Granger causality for inspecting directed interdependencies of neural assemblies," *Journal of Neuroscience Methods*, vol. 124, no. 1, pp. 27-44, 2003.

- [Hiemstra 1994] Hiemstra, C., and Jones, J. D., "Testing for linear and nonlinear Granger causality in the stock price-volume relation," *Journal of Finance*, vol. 49, no. 5, pp. 1639-1664, 1994.
- [Hirtz 2007] Hirtz, D., Thurman, D. J., Gwinn-Hardy, K., Mohamed, M., Chaudhuri, A. R., and Zalutsky, R., "How common are the "common" neurologic disorders?," *Neurology*, vol. 68, no. 5, pp. 326-337, 2007.
- [Honey 2007] Honey, C. J., Kötter, R., Breakspear, M., and Sporns, O., "Network structure of cerebral cortex shapes functional connectivity on multiple time scales," *Proceedings of the National Academy of Sciences*, vol. 104, no. 24, pp. 10240-10245, 2007.
- [Horwitz 2003] Horwitz, B., "The elusive concept of brain connectivity," *NeuroImage*, vol. 19, no. 2, pp. 466-470, 2003.
- [Hosoya 2001] Hosoya, Y., "Elimination of third-series effect and defining partial measures of causality," *Journal of Time Series Analysis*, vol. 22, no. 5, pp. 537-554, 2001.
- [Hubel 1963] Hubel, D. H., and Wiesel, T. N., "Shape and arrangement of columns in cat's striate cortex," *The Journal of Physiology*, vol. 165, no. 3, pp. 559-568, 1963.
- [Hubel 1965] Hubel, D. H., and Wiesel, T. N., "Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat," *Journal of Neurophysiology*, vol. 28, no. 2, pp. 229-289, 1965.
- [Hubel 1968] Hubel, D. H., and Wiesel, T. N., "Receptive fields and functional architecture of monkey striate cortex," *The Journal of Physiology*, vol. 195, no. 1, pp. 215-243, 1968.
- [Iläe 1993] Iläe, "Guidelines for epidemiologic studies on epilepsy," *Epilepsia*, vol. 34, no. 4, pp. 592-596, 1993.
- [Jansen 1993] Jansen, B. H., Zouridakis, G., and Brandt, M. E., "A neurophysiologically-based mathematical model of flash visual evoked potentials," *Biological Cybernetics*, vol. 68, no. 3, pp. 275-283, 1993.
- [Jansen 1995] Jansen, B. H., and Rit, V. G., "Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns," *Biological Cybernetics*, vol. 73, no. 4, pp. 357-366, 1995.
- [Jilek-Aall 1999] Jilek-Aall, L., "Morbus sacer in Africa: some religious aspects of epilepsy in traditional cultures," *Epilepsia*, vol. 40, no. 3, pp. 382-386, 1999.
- [Johansen-Berg 2006] Johansen-Berg, H., and Behrens, T. E. J., "Just pretty pictures? What diffusion tractography can add in clinical neuroscience," *Current Opinion in Neurology*, vol. 19, no. 4, pp. 379, 2006.
- [Johansen-Berg 2009] Johansen-Berg, H., and Rushworth, M. F. S., "Using diffusion imaging to study human connective anatomy," *Annual Review of Neuroscience*, vol. 32, pp. 75-94, 2009.
- [Jung 2012] Jung, Y. J., Kim, K. H., and Im, C. H., "Mathematical Issues in the Inference of Causal Interactions among Multichannel Neural Signals," *Journal of Applied Mathematics*, vol. 2012, pp. 1-14, 2012.

- [Kamiński 1991] Kamiński, M. J., and Blinowska, K. J., "A new method of the description of the information flow in the brain structures," *Biological Cybernetics*, vol. 65, no. 3, pp. 203-210, 1991.
- [Kamiński 1997] Kamiński, M. J., Blinowska, K. J., and Szelenberger, W., "Topographic analysis of coherence and propagation of EEG activity during sleep and wakefulness," *Electroencephalography and Clinical Neurophysiology*, vol. 102, no. 3, pp. 216-227, 1997.
- [Kamiński 2001] Kamiński, M. J., Ding, M., Truccolo, W. A., and Bressler, S. L., "Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance," *Biological Cybernetics*, vol. 85, no. 2, pp. 145-157, 2001.
- [Kamiński 2005] Kamiński, M. J., "Determination of transmission patterns in multichannel data," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 360, no. 1457, pp. 947-952, 2005.
- [Kantz 1997] Kantz, H., and Schreiber, T., *Nonlinear Time Series Analysis*, 2nd ed., Cambridge, MA: Cambridge university press Cambridge, 1997.
- [Kariya 2004] Kariya, T., and Kurata, H., *Generalized Least Squares*, Chichester: John Wiley & Sons, Ltd, 2004.
- [Katz 1981] Katz, R. W., "On some criteria for estimating the order of a Markov chain," *Technometrics*, vol. 23, no. 3, pp. 243-249, 1981.
- [Kiebel 2009] Kiebel, S. J., Garrido, M. I., Moran, R., Chen, C. C., and Friston, K. J., "Dynamic causal modeling for EEG and MEG," *Human Brain Mapping*, vol. 30, no. 6, pp. 1866-1876, 2009.
- [Kleeman 2011] Kleeman, R., "Information Theory and Dynamical System Predictability," *Entropy*, vol. 13, no. 3, pp. 612-649, 2011.
- [Klingler 1956] Klingler, J., and Ludwig, E., *Atlas Cerebri Humani*, 1 ed.: Basel, 1956.
- [Knobler 2004] Knobler, S., *The infectious etiology of chronic diseases: defining the relationship, enhancing the research, and mitigating the effects*, Washington: The National Academies Press, 2004.
- [Knyazeva 2010] Knyazeva, M. G., Jalili, M., Brioschi, A., Bourquin, I., Fornari, E., Hasler, M., Meuli, R., Maeder, P., and Ghika, J., "Topography of EEG multivariate phase synchronization in early Alzheimer's disease," *Neurobiology of Aging*, vol. 31, no. 7, pp. 1132-1144, 2010.
- [Kocsis 2006] Kocsis, B., and Kaminski, M., "Dynamic changes in the direction of the theta rhythmic drive between supramammillary nucleus and the septohippocampal system," *Hippocampus*, vol. 16, no. 6, pp. 531-540, 2006.
- [Korzeniewska 1997] Korzeniewska, A., Kasicki, S., Kamiński, M. J., and Blinowska, K. J., "Information flow between hippocampus and related structures during various types of rat's behavior," *Journal of Neuroscience Methods*, vol. 73, no. 1, pp. 49-60, 1997.
- [Kramer 2005] Kramer, M. A., Kirsch, H. E., and Szeri, A. J., "Pathological pattern formation and cortical propagation of epileptic seizures," *Journal of the Royal Society Interface*, vol. 2, no. 2, pp. 113-127, 2005.

- [Kramer 2006] Kramer, M. A., Lopour, B. A., Kirsch, H. E., and Szeri, A. J., "Bifurcation control of a seizing human cortex," *Physical Review E*, vol. 73, no. 4, pp. 041928-1:041928-16, 2006.
- [Kramer 2007] Kramer, M. A., Chang, F. L., Cohen, M. E., Hudson, D., and Szeri, A. J., "Synchronization measures of the scalp electroencephalogram can discriminate healthy from Alzheimer's subjects," *International Journal of Neural Systems*, vol. 17, no. 2, pp. 61-70, 2007.
- [Kullback 1997] Kullback, S., *Information Theory and Statistics*, New York: Dover Pubns, 1997.
- [Kwan 2000] Kwan, P., and Brodie, M. J., "Early identification of refractory epilepsy," *New England Journal of Medicine*, vol. 342, no. 5, pp. 314-319, 2000.
- [Lachaux 1999] Lachaux, J. P., Rodriguez, E., Martinerie, J., and Varela, F. J., "Measuring phase synchrony in brain signals," *Human Brain Mapping*, vol. 8, no. 4, pp. 194-208, 1999.
- [Le Bihan 1986] Le Bihan, D., Breton, E., Lallemand, D., Grenier, P., Cabanis, E., and Laval-Jeantet, M., "MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders," *Radiology*, vol. 161, no. 2, pp. 401-407, 1986.
- [Levisohn 2007] Levisohn, P. M., "The autism-epilepsy connection," *Epilepsia*, vol. 48, no. 9, pp. 33-35, 2007.
- [Liégeois 2004] Liégeois, F., Connelly, A., Cross, J. H., Boyd, S. G., Gadian, D. G., Vargha-Khadem, F., and Baldeweg, T., "Language reorganization in children with early - onset lesions of the left hemisphere: an fMRI study," *Brain*, vol. 127, no. 6, pp. 1229-1236, 2004.
- [Lindner 2011] Lindner, M., Vicente, R., Priesemann, V., and Wibral, M., "TRENTOOL: A Matlab open source toolbox to analyse information flow in time series data with transfer entropy," *BMC Neuroscience*, vol. 12, no. 1, pp. 119, 2011.
- [Lopes Da Silva 1974] Lopes Da Silva, F. H., Hoeks, A., Smits, H., and Zetterberg, L. H., "Model of brain rhythmic activity," *Biological Cybernetics*, vol. 15, no. 1, pp. 27-37, 1974.
- [Lopes Da Silva 1976] Lopes Da Silva, F. H., Van Rotterdam, A., Barts, P., Van Heusden, E., and Burr, W., *Model of Neuronal Populations - The Basic Mechanism of Rhythmicity*, Amsterdam: Elsevier, 1976.
- [Lopes Da Silva 1989] Lopes Da Silva, F. H., Pijn, J. P., and Boeijinga, P., "Interdependence of EEG signals: linear vs. nonlinear associations and the significance of time delays and phase shifts," *Brain Topography*, vol. 2, no. 1, pp. 9-18, 1989.
- [Lopes Da Silva 2003] Lopes Da Silva, F. H., Blanes, W., Kalitzin, S. N., Parra, J., Suffczynski, P., and Velis, D. N., "Dynamical diseases of brain systems: different routes to epileptic seizures," *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 5, pp. 540-548, 2003.
- [Lorenz 1963] Lorenz, E. N., "Deterministic Nonperiodic Flow," *Journal of the Atmospheric Sciences*, vol. 20, pp. 130-141, 1963.
- [Loring 2004] Loring, D. W., Meador, K. J., and Lee, G. P., "Determinants of quality of life in epilepsy," *Epilepsy & Behavior*, vol. 5, no. 6, pp. 976-980, 2004.
- [Louis Dorr 2007] Louis Dorr, V., Caparos, M., Wendling, F., Vignal, J. P., and Wolf, D., "Extraction of reproducible seizure patterns based on EEG scalp correlations," *Biomedical Signal Processing and Control*, vol. 2, no. 3, pp. 154-162, 2007.

- [Mac 2007] Mac, T. L., Tran, D. S., Quet, F., Odermatt, P., Preux, P. M., and Tan, C. T., "Epidemiology, aetiology, and clinical management of epilepsy in Asia: a systematic review," *The Lancet Neurology*, vol. 6, no. 6, pp. 533-543, 2007.
- [Macdonald 2000] Macdonald, B. K., Cockerell, O. C., Sander, J., and Shorvon, S. D., "The incidence and lifetime prevalence of neurological disorders in a prospective community-based study in the UK," *Brain*, vol. 123, no. 4, pp. 665-676, 2000.
- [Mangin 1998] Mangin, J. F., Régis, J., Poline, J. B., Rivière, D., Poupon, C., Poupon, F., Papadopoulos, D., Delaye, F., Pizzato, O., and Coulon, O., "Place de l'anatomie dans la cartographie fonctionnelle du cerveau," *Annales de l'Institut Pasteur*, vol. 9, no. 3, pp. 243-258, 1998.
- [Marinazzo 2008] Marinazzo, D., Pellicoro, M., and Stramaglia, S., "Kernel method for nonlinear Granger causality," *Physical Review Letters*, vol. 100, no. 14, pp. 144103, 2008.
- [Marson 1996] Marson, A. G., Kadir, Z. A., and Chadwick, D. W., "New antiepileptic drugs: a systematic review of their efficacy and tolerability," *British Medical Journal*, vol. 313, no. 7066, pp. 1169-1174, 1996.
- [Martinerie 1992] Martinerie, J. M., Albano, A. M., Mees, A. I., and Rapp, P. E., "Mutual information, strange attractors, and the optimal estimation of dimension," *Physical Review A*, vol. 45, no. 10, pp. 7058-7064, 1992.
- [Martini 2011] Martini, M., Kranz, T. A., Wagner, T., and Lehnertz, K., "Inferring directional interactions from transient signals with symbolic transfer entropy," *Physical Review E*, vol. 83, no. 1, pp. 011919, 2011.
- [Massimini 2005] Massimini, M., Ferrarelli, F., Huber, R., Esser, S. K., Singh, H., and Tononi, G., "Breakdown of cortical effective connectivity during sleep," *Science*, vol. 309, no. 5744, pp. 2228, 2005.
- [Mckee 2010] Mckee, R. "Partial (focal) seizure," http://ehealthforum.com/blogs/colored_stone/partial-focal-seizure-b11015.html.
- [McIntosh 1994] McIntosh, A. R., and Gonzalez-Lima, F., "Structural equation modeling and its application to network analysis in functional brain imaging," *Human Brain Mapping*, vol. 2, no. 1-2, pp. 2-22, 1994.
- [Medvedev 1999] Medvedev, A., and Willoughby, J. O., "Autoregressive modeling of the EEG in systemic kainic acid-induced epileptogenesis," *International Journal of Neuroscience*, vol. 97, no. 3-4, pp. 149-167, 1999.
- [Meunier 2009] Meunier, D., Achard, S., Morcom, A., and Bullmore, E., "Age-related changes in modular organization of human brain functional networks," *NeuroImage*, vol. 44, no. 3, pp. 715-723, 2009.
- [Mitra 2008] Mitra, P., and Bokil, H., *Observed Brain Dynamics*: Oxford University Press, USA, 2008.
- [Moran 2007] Moran, R. J., Kiebel, S. J., Stephan, K. E., Reilly, R. B., Daunizeau, J., and Friston, K. J., "A neural mass model of spectral responses in electrophysiology," *NeuroImage*, vol. 37, no. 3, pp. 706-720, 2007.
- [Morris 1981] Morris, C., and Lecar, H., "Voltage oscillations in the barnacle giant muscle fiber," *Biophysical Journal*, vol. 35, no. 1, pp. 193-213, 1981.

- [Mountcastle 1957] Mountcastle, V. B., "Modality and topographic properties of single neurons of cat's somatic sensory cortex," *Journal of Neurophysiology*, vol. 20, no. 4, pp. 408-434, 1957.
- [Muñoz 2007] Muñoz, A., Méndez, P., Defelipe, J., and Alvarez-Leefmans, F. J., "Cation-chloride cotransporters and GABA-ergic innervation in the human epileptic hippocampus," *Epilepsia*, vol. 48, no. 4, pp. 663-673, 2007.
- [Neymotin 2011] Neymotin, S. A., Jacobs, K. M., Fenton, A. A., and Lytton, W. W., "Synaptic information transfer in computer models of neocortical columns," *Journal of Computational Neuroscience*, vol. 30, no. 1, pp. 69-84, 2011.
- [Niedermeyer 1999] Niedermeyer, E., and Lopes Da Silva, F., *Electroencephalography: Basic Principles*, 4th ed., Baltimore: Williams & Wilkins, 1999.
- [Nishiyama 1997] Nishiyama, Y., "Exports' contribution to economic growth: Empirical evidence for California, Massachusetts, and Texas, using employment data," *Journal of Regional Science*, vol. 37, no. 1, pp. 99-125, 1997.
- [Noebels 1996] Noebels, J. L., "Targeting Epilepsy Genes Minireview," *Neuron*, vol. 16, no. 2, pp. 241-244, 1996.
- [Nolte 2008] Nolte, G., Ziehe, A., Nikulin, V. V., Schlögl, A., Krämer, N., Brismar, T., and Müller, K. R., "Robustly estimating the flow direction of information in complex physical systems," *Physical Review Letters*, vol. 100, no. 23, pp. 234101, June, 2008.
- [Nolte 2009] Nolte, G., Ziehe, A., Krämer, N., Popescu, F., and Müller, K. R., "Comparison of Granger causality and phase slope index," pp. 1-10, 2009.
- [Nolte 2010] Nolte, G., and Müller, K. R., "Localizing and estimating causal relations of interacting brain rhythms," *Frontiers in Human Neuroscience*, vol. 4, no. 209, pp. 1-5, 2010.
- [Nunez 1995] Nunez, P. L., *Neocortical Dynamics and Human EEG Rhythms*: Oxford University Press, USA, 1995.
- [Olson 2008] Olson, D. "Differentiating epileptic seizures from nonepileptic spells," <http://www.pediatricsconsultant360.com/content/differentiating-epileptic-seizures-nonepileptic-spells>.
- [Ooyen 2001] Ooyen, A. V., "Competition in the development of nerve connections: a review of models," *Network: Computation in Neural Systems*, vol. 12, no. 1, pp. 1-47, 2001.
- [Paus 1997] Paus, T., Jech, R., Thompson, C. J., Comeau, R., Peters, T., and Evans, A. C., "Transcranial magnetic stimulation during positron emission tomography: a new method for studying connectivity of the human cerebral cortex," *Journal of Neuroscience*, vol. 17, no. 9, pp. 3178-3184, 1997.
- [Pearl 2000] Pearl, J., *Causality: models, reasoning and inference*, Cambridge, U.K.: Cambridge University Press, 2000.
- [Pecora 1990] Pecora, L. M., and Carroll, T. L., "Synchronization in chaotic systems," *Physical Review Letters*, vol. 64, no. 8, pp. 821-824, 1990.
- [Penny 2010] Penny, W. D., Stephan, K. E., Daunizeau, J., Rosa, M. J., Friston, K. J., Schofield, T. M., and Leff, A. P., "Comparing families of dynamic causal models," *PLoS Computational Biology*, vol. 6, no. 3, pp. e1000709, 2010.

- [Pereda 2005] Pereda, E., Quiroga, R. Q., and Bhattacharya, J., "Nonlinear multivariate analysis of neurophysiological signals," *Progress in Neurobiology*, vol. 77, no. 1-2, pp. 1-37, 2005.
- [Pfurtscheller 1999a] Pfurtscheller, G., and Andrew, C., "Event-related changes of band power and coherence: methodology and interpretation," *Journal of Clinical Neurophysiology*, vol. 16, no. 6, pp. 512-519, 1999a.
- [Pfurtscheller 1999b] Pfurtscheller, G., and Lopes Da Silva, F. H., "Event-related EEG/MEG synchronization and desynchronization: basic principles," *Clinical Neurophysiology*, vol. 110, no. 11, pp. 1842-1857, 1999b.
- [Piersol 1993] Piersol, A. G., and Bendat, J. S., *Engineering Applications of Correlation and Spectral Analysis*, 2nd ed., New York: Wiley, 1993.
- [Pikovsky 1984] Pikovsky, A. S., "On the interaction of strange attractors," *Zeitschrift Für Physik B Condensed Matter*, vol. 55, no. 2, pp. 149-154, 1984.
- [Pikovsky 2003a] Pikovsky, A. S., Rosenblum, M., and Kurths, J., *Synchronization: Theory and Application*: Springer Netherlands, 2003a.
- [Pikovsky 2003b] Pikovsky, A. S., Rosenblum, M., and Kurths, J., *Synchronization: A Universal Concept in Nonlinear Sciences*: Cambridge University Press, 2003b.
- [Popescu 2011] Popescu, F., "Robust Statistics for Describing Causality in Multivariate Time Series," *Journal of Machine Learning*, vol. 12, pp. 30-64, 2011.
- [Quiroga 2000] Quiroga, R. Q., Arnhold, J., and Grassberger, P., "Learning driver-response relationships from synchronization patterns," *Physical Review E*, vol. 61, no. 5, pp. 5142, 2000.
- [Rana 2012] Rana, P., Lipor, J., Lee, H., Van Drongelen, W., Kohrman, M. H., and Vanveen, B., "Seizure detection using the Phase-Slope Index and multichannel ECoG," *IEEE Transactions on Biomedical Engineering*, no. 99, pp. 1-1, 2012.
- [Rosenblum 2004] Rosenblum, M. G., Pikovsky, A. S., and Kurths, J., "Synchronization approach to analysis of biological systems," *Fluctuation and Noise Lett*, vol. 4, pp. L53-L62, 2004.
- [Rubinov 2010] Rubinov, M., and Sporns, O., "Complex network measures of brain connectivity: uses and interpretations," *NeuroImage*, vol. 52, no. 3, pp. 1059-1069, 2010.
- [Rudrauf 2006] Rudrauf, D., Douiri, A., Kovach, C., Lachaux, J. P., Cosmelli, D., Chávez, M., Adam, C., Renault, B., Martinerie, J., and Le Van Quyen, M., "Frequency flows and the time-frequency dynamics of multivariate phase synchronization in brain signals," *NeuroImage*, vol. 31, no. 1, pp. 209-227, 2006.
- [Ruff 2009] Ruff, C. C., Driver, J., and Bestmann, S., "Combining TMS and fMRI: From 'virtual lesions' to functional-network accounts of cognition," *Cortex*, vol. 45, no. 9, pp. 1043-1049, 2009.
- [Sabesan 2003] Sabesan, S., Narayanan, K., Prasad, A., Spanias, A., and Iasemidis, L. D., "Improved measure of information flow in coupled nonlinear systems," *Proceedings of the IASTED International Conference, Modelling and Simulation, California, US*, pp. 329-333, 2003.

- [Sabesan 2007] Sabesan, S., Narayanan, K., Prasad, A., Iasemidis, L. D., Spanias, A., and Tsakalis, K., *Information flow in coupled nonlinear systems: application to the epileptic human brain*, New York: Springer, 2007.
- [Sabesan 2009a] Sabesan, S., Chakravarthy, N., Tsakalis, K., Pardalos, P., and Iasemidis, L., "Measuring resetting of brain dynamics at epileptic seizures: application of global optimization and spatial synchronization techniques," *Journal of Combinatorial Optimization*, vol. 17, no. 1, pp. 74-97, 2009a.
- [Sabesan 2009b] Sabesan, S., Good, L., Tsakalis, K., Spanias, A., Treiman, D., and Iasemidis, L., "Information flow and application to epileptogenic focus localization from intracranial EEG," *IEEE Transactions on Neural Systems and Rehabilitation Engineering: a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 17, no. 3, pp. 244-253, June, 2009b.
- [Saito 1981] Saito, Y., and Harashima, H., *Tracking of information within multichannel EEG record-causal analysis in EEG*, Amsterdam: Elsevier-North-Holland Biomedical Press, 1981.
- [Sakkalis 2009] Sakkalis, V., Giurcaneanu, C. D., Xanthopoulos, P., Zervakis, M. E., Tsiaras, V., Yang, Y., Karakonstantaki, E., and Micheloyannis, S., "Assessment of linear and nonlinear synchronization measures for analyzing EEG in a mild epileptic paradigm," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 4, pp. 433-441, 2009.
- [Sakkalis 2011] Sakkalis, V., "Review of advanced techniques for the estimation of brain connectivity measured with EEG/MEG," *Computers in Biology and Medicine*, vol. 41, no. 12, pp. 1110-1117, 2011.
- [Salazar 2004] Salazar, R. F., König, P., and Kayser, C., "Directed interactions between visual areas and their role in processing image structure and expectancy," *European Journal of Neuroscience*, vol. 20, no. 5, pp. 1391-1401, 2004.
- [Sameshima 1999] Sameshima, K., and Baccala, L. A., "Using partial directed coherence to describe neuronal ensemble interactions," *Journal of Neuroscience Methods*, vol. 94, no. 1, pp. 93-103, 1999.
- [Sander 2003] Sander, J. W., "The epidemiology of epilepsy revisited," *Current Opinion in Neurology*, vol. 16, no. 2, pp. 165-170, 2003.
- [Schelter 2006a] Schelter, B., Winterhalder, M., Eichler, M., Peifer, M., Hellwig, B., Guschlbauer, B., Lüking, C. H., Dahlhaus, R., and Timmer, J., "Testing for directed influences among neural signals using partial directed coherence," *Journal of Neuroscience Methods*, vol. 152, no. 1-2, pp. 210-219, 2006a.
- [Schelter 2006b] Schelter, B., Winterhalder, M., Hellwig, B., Guschlbauer, B., Lüking, C. H., and Timmer, J., "Direct or indirect? Graphical models for neural oscillators," *Journal of Physiology-Paris*, vol. 99, no. 1, pp. 37-46, 2006b.
- [Scherg 1985a] Scherg, M., and Von Cramon, D., "A new interpretation of the generators of BAEP waves IV: results of a spatio-temporal dipole model," *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, vol. 62, no. 4, pp. 290-299, 1985a.
- [Scherg 1985b] Scherg, M., and Von Cramon, D., "Two bilateral sources of the late AEP as identified by a spatio-temporal dipole model," *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, vol. 62, no. 1, pp. 32-44, 1985b.

- [Schnider 1989] Schnider, S. M., Kwong, R. H., Lenz, F. A., and Kwan, H. C., "Detection of feedback in the central nervous system using system identification techniques," *Biological Cybernetics*, vol. 60, no. 3, pp. 203-212, 1989.
- [Schnupp 2006] Schnupp, J. W. H., Hall, T. M., Kokelaar, R. F., and Ahmed, B., "Plasticity of temporal pattern codes for vocalization stimuli in primary auditory cortex," *Journal of Neuroscience*, vol. 26, no. 18, pp. 4785-4795, 2006.
- [Schreiber 2000] Schreiber, T., "Measuring information transfer," *Physical Review Letters*, vol. 85, no. 2, pp. 461-464, 2000.
- [Schwarz 1978] Schwarz, G., "Estimating the dimension of a model," *The Annals of Statistics*, vol. 6, no. 2, pp. 461-464, 1978.
- [Sekihara 2010] Sekihara, K., Owen, J., Attias, H., Wipf, D., and Nagarajan, S. S., "Estimating directions of information flow between cortical activities using phase-slope index," *17th International Conference on Biomagnetism Advances in Biomagnetism - BIOMAG2010, IFMBE Proceedings*, pp. 199-202, 2010.
- [Senhadji 2009] Senhadji, L., Ansari-Asl, K., and Wendling, F., "From EEG signals to brain connectivity: methods and applications in epilepsy," *Advanced Biosignal Processing*, pp. 145-164, 2009.
- [Seth 2005] Seth, A. K., Baars, B. J., and Edelman, D. B., "Criteria for consciousness in humans and other mammals," *Consciousness and Cognition*, vol. 14, no. 1, pp. 119-139, 2005.
- [Seth 2007] Seth, A. K., and Edelman, G. M., "Distinguishing causal interactions in neural populations," *Neural Computation*, vol. 19, no. 4, pp. 910-933, 2007.
- [Shannon 1949] Shannon, C. E., and Weaver, W., *The Mathematical Theory of Communication*, Urbana, Illinois: University of Illinois Press, 1949.
- [Sheth 2009] Sheth, R. A., Josephson, L., and Mahmood, U., "Evaluation and clinically relevant applications of a fluorescent imaging analog to fluorodeoxyglucose positron emission tomography," *Journal of Biomedical Optics*, vol. 14, no. 6, pp. 064014, 2009.
- [Shu 2002] Shu, H. Z., Luo, L. M., Zhou, J. D., and Bao, X. D., "Moment-based methods for polygonal approximation of digitized curves," *Pattern Recognition*, vol. 35, no. 2, pp. 421-434, 2002.
- [Skinner 2005a] Skinner, F. K., Bazzazi, H., and Campbell, S. A., "Two-cell to N-cell heterogeneous, inhibitory networks: Precise linking of multistable and coherent properties," *Journal of Computational Neuroscience*, vol. 18, no. 3, pp. 343-352, 2005a.
- [Skinner 2005b] Skinner, F. K., Chung, J. Y. J., Ncube, I., Murray, P. A., and Campbell, S. A., "Using heterogeneity to predict inhibitory network model characteristics," *Journal of Neurophysiology*, vol. 93, no. 4, pp. 1898-1907, 2005b.
- [Sperling 1999] Sperling, M. R., Feldman, H., Kinman, J., Liporace, J. D., and Oconnor, M. J., "Seizure control and mortality in epilepsy," *Annals of Neurology*, vol. 46, no. 1, pp. 45-50, 1999.
- [Sporns 2000] Sporns, O., Tononi, G., and Edelman, G. M., "Connectivity and complexity: the relationship between neuroanatomy and brain dynamics," *Neural Networks*, vol. 13, no. 8-9, pp. 909-922, 2000.

- [Sporns 2004] Sporns, O., Chialvo, D. R., Kaiser, M., and Hilgetag, C. C., "Organization, development and function of complex brain networks," *Trends in Cognitive Sciences*, vol. 8, no. 9, pp. 418-425, 2004.
- [Sporns 2005] Sporns, O., Tononi, G., and Kötter, R., "The human connectome: a structural description of the human brain," *PLoS Computational Biology*, vol. 1, no. 4, pp. 245-251, 2005.
- [Sporns 2007] Sporns, O., 2007, http://www.scholarpedia.org/article/Brain_connectivity.
- [Sporns 2010] Sporns, O., *Networks of the Brain*, London, England: The MIT Press, 2010.
- [Stam 2002a] Stam, C. J., Van Cappellen Van Walsum, A. M., Pijnenburg, Y. A. L., Berendse, H. W., De Munck, J. C., Scheltens, P., and Van Dijk, B. W., "Generalized synchronization of MEG recordings in Alzheimer's disease: Evidence for involvement of the gamma band," *Journal of Clinical Neurophysiology*, vol. 19, no. 6, pp. 562, 2002a.
- [Stam 2002b] Stam, C. J., and Van Dijk, B. W., "Synchronization likelihood: an unbiased measure of generalized synchronization in multivariate data sets," *Physica D: Nonlinear Phenomena*, vol. 163, no. 3-4, pp. 236-251, 2002b.
- [Stephan 2008] Stephan, K. E., Kasper, L., Harrison, L. M., Daunizeau, J., Den Ouden, H. E. M., Breakspear, M., and Friston, K. J., "Nonlinear dynamic causal models for fMRI," *NeuroImage*, vol. 42, no. 2, pp. 649-662, 2008.
- [Suffczynski 2001] Suffczynski, P., Kalitzin, S., Pfurtscheller, G., and Lopes Da Silva, F. H., "Computational model of thalamo-cortical networks: dynamical control of alpha rhythms in relation to focal attention," *International Journal of Psychophysiology*, vol. 43, no. 1, pp. 25-40, 2001.
- [Suffczynski 2004] Suffczynski, P., Kalitzin, S., and Lopes Da Silva, F. H., "Dynamics of non-convulsive epileptic phenomena modeled by a bistable neuronal network," *Neuroscience*, vol. 126, no. 2, pp. 467-484, 2004.
- [Suffczynski 2006] Suffczynski, P., Lopes Da Silva, F. H., Parra, J., Velis, D. N., Bouwman, B. M., Van Rijn, C. M., Van Hese, P., Boon, P., Khosravani, H., and Derchansky, M., "Dynamics of epileptic phenomena determined from statistics of ictal transitions," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 3, pp. 524-532, 2006.
- [T.N.S.E. 2009] T.N.S.E., "What is Epilepsy?," [The National Society for Epilepsy, 2009], Publisher, <http://www.epilepsysociety.org.uk/AboutEpilepsy/Whatisepilepsy>.
- [Takigawa 1988] Takigawa, M., "Rhythmic light therapy for depression and data processing analysis of its effects by directed coherence," *Activitas Nervosa Superior*, vol. 30, no. 3, pp. 177-180, 1988.
- [Takigawa 1996] Takigawa, M., Wang, G., Kawasaki, H., and Fukuzako, H., "EEG analysis of epilepsy by directed coherence method a data processing approach," *International Journal of Psychophysiology*, vol. 21, no. 2-3, pp. 65-73, 1996.
- [Theiler 1992] Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., and Doyne Farmer, J., "Testing for nonlinearity in time series: the method of surrogate data," *Physica D: Nonlinear Phenomena*, vol. 58, no. 1-4, pp. 77-94, 1992.
- [Tomson 2005] Tomson, T., Walczak, T., Sillanpaa, M., and Sander, J. W. A. S., "Sudden unexpected death in epilepsy: a review of incidence and risk factors," *Epilepsia*, vol. 46, pp. 54-61, 2005.

- [Tong 1975] Tong, H., "Determination of the order of a Markov chain by Akaike's information criterion," *Journal of Applied Probability*, vol. 12, no. 3, pp. 488-497, 1975.
- [Tononi 1994] Tononi, G., Sporns, O., and Edelman, G. M., "A measure for brain complexity: relating functional segregation and integration in the nervous system," *Proceedings of the National Academy of Sciences*, vol. 91, no. 11, pp. 5033-5037, 1994.
- [Tononi 2003] Tononi, G., and Sporns, O., "Measuring information integration," *BMC Neuroscience*, vol. 4, no. 1, pp. 31-50, 2003.
- [Touboul 2011] Touboul, J., Wendling, F., Chauvel, P., and Faugeras, O., "Neural mass activity, bifurcations, and epilepsy," *Neural Computation*, vol. 23, no. 12, pp. 3232-3286, 2011.
- [Tramonte 2008] Tramonte, S. M., *Generalized Tonic-Clonic Seizures*, 7th ed., New York: McGraw-Hill, 2008.
- [Traub 1982] Traub, R. D., and Wong, R. K., "Cellular mechanism of neuronal synchronization in epilepsy," *Science*, vol. 216, no. 4547, pp. 745-747, 1982.
- [Traub 2001] Traub, R. D., Whittington, M. A., Buhl, E. H., Lebeau, F. E. N., Bibbig, A., Boyd, S., Cross, H., and Baldeweg, T., "A possible role for gap junctions in generation of very fast EEG oscillations preceding the onset of, and perhaps initiating, seizures," *Epilepsia*, vol. 42, no. 2, pp. 153-170, 2001.
- [Traub 2003] Traub, R. D., Pais, I., Bibbig, A., Lebeau, F. E. N., Buhl, E. H., Hormuzdi, S. G., Monyer, H., and Whittington, M. A., "Contrasting roles of axonal (pyramidal cell) and dendritic (interneuron) electrical coupling in the generation of neuronal network oscillations," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 3, pp. 1370-1374, 2003.
- [Traub 2005] Traub, R. D., Contreras, D., Cunningham, M. O., Murray, H., Lebeau, F. E. N., Roopun, A., Bibbig, A., Wilent, W. B., Higley, M. J., and Whittington, M. A., "Single-column thalamocortical network model exhibiting gamma oscillations, sleep spindles, and epileptogenic bursts," *Journal of Neurophysiology*, vol. 93, no. 4, pp. 2194-2232, 2005.
- [Treserras 2008] Treserras, S., "Etudes sur la connectivité cérébrale: aspects méthodologiques et applications au cerveau au repos, à la motricité et à la lecture," *Imagerie Médicale et Handicaps Neurologiques*, INSERM UMR 825, Université de Toulouse, Université Toulouse III-Paul Sabatier, Toulouse, 2008.
- [Trevelyan 2006] Trevelyan, A. J., Sussillo, D., Watson, B. O., and Yuste, R., "Modular propagation of epileptiform activity: evidence for an inhibitory veto in neocortex," *The Journal of Neuroscience*, vol. 26, no. 48, pp. 12447-12455, 2006.
- [Uva 2005] Uva, L., Librizzi, L., Wendling, F., and De Curtis, M., "Propagation dynamics of epileptiform activity acutely induced by Bicuculline in the hippocampal-parahippocampal region of the isolated guinea pig brain," *Epilepsia*, vol. 46, no. 12, pp. 1914-1925, 2005.
- [Vakorin 2009] Vakorin, V. A., Krakovska, O. A., and McIntosh, A. R., "Confounding effects of indirect connections on causality estimation," *Journal of Neuroscience Methods*, vol. 184, no. 1, pp. 152-160, 2009.
- [Valdes-Sosa 2011] Valdes-Sosa, P. A., Roebroeck, A., Daunizeau, J., and Friston, K., "Effective connectivity: Influence, causality and biophysical modeling," *NeuroImage*, vol. 58, no. 2, pp. 339-361, 2011.

- [Valencia 2008] Valencia, M., Martinerie, J., Dupont, S., and Chávez, M., "Dynamic small-world behavior in functional brain networks unveiled by an event-related networks approach," *Physical Review E*, vol. 77, no. 5, pp. 050905-1:050905-4, 2008.
- [Van Drongelen 2005] Van Drongelen, W., Lee, H. C., Hereld, M., Chen, Z., Elsen, F. P., and Stevens, R. L., "Emergent epileptiform activity in neural networks with weak excitatory synapses," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 13, no. 2, pp. 236-241, 2005.
- [Van Drongelen 2007] Van Drongelen, W., Lee, H. C., Stevens, R. L., and Hereld, M., "Propagation of seizure-like activity in a model of neocortex," *Journal of Clinical Neurophysiology*, vol. 24, no. 2, pp. 182-188, 2007.
- [Veeramani 2004] Veeramani, B., Narayanan, K., Prasad, A., Iasemidis, L. D., Spanias, A. S., and Tsakalis, K., "Measuring the direction and the strength of coupling in nonlinear systems-a modeling approach in the state space," *IEEE Signal Processing Letters* vol. 11, no. 7, pp. 617-620, 2004.
- [Vélez-Pérez 2008] Vélez-Pérez, H., Louis Dorr, V., Ranta, R., and Dufaut, M., "Connectivity estimation of three parametric methods on simulated electroencephalogram signals," *30th Annual Conference of the IEEE EMB Society, Vancouver, Canada*. pp. 2606-2609, 2008.
- [Vicente 2011] Vicente, R., Wibral, M., Lindner, M., and Pipa, G., "Transfer entropy - a model-free measure of effective connectivity for the neurosciences," *Journal of Computational Neuroscience*, vol. 30, pp. 45-67, 2011.
- [Vieussens 1684] Vieussens, R., *Nevrographia Universalis*, 1st ed., Lyon: Apud Joannem Certe, 1684.
- [W.H.O. 2001] W.H.O., 2001, http://allcountries.org/health/epilepsy_historical_overview.html.
- [Wang 2007] Wang, X., Chen, Y., Bressler, S. L., and Ding, M., "Granger causality between multiple interdependent neurobiological time series: Blockwise versus pairwise methods," *International Journal of Neural Systems*, vol. 17, no. 2, pp. 71-78, 2007.
- [Website1 2001] Website1, 2001, <http://www.highlands.edu/academics/divisions/scipe/biology/faculty/harnden/2121/images/fibersnuclei.jpg>.
- [Wen 2005] Wen, Q., and Chklovskii, D. B., "Segregation of the brain into gray and white matter: a design minimizing conduction delays," *PLoS Computational Biology*, vol. 1, no. 7, pp. e78, 2005.
- [Wendling 2000] Wendling, F., Bellanger, J. J., Bartolomei, F., and Chauvel, P., "Relevance of nonlinear lumped-parameter models in the analysis of depth-EEG epileptic signals," *Biological Cybernetics*, vol. 83, no. 4, pp. 367-378, 2000.
- [Wendling 2001] Wendling, F., Bartolomei, F., Bellanger, J. J., and Chauvel, P., "Interpretation of interdependencies in epileptic signals using a macroscopic physiological model of the EEG," *Clinical Neurophysiology*, vol. 112, no. 7, pp. 1201-1218, 2001.
- [Wendling 2002] Wendling, F., Bartolomei, F., Bellanger, J. J., and Chauvel, P., "Epileptic fast activity can be explained by a model of impaired GABAergic dendritic inhibition," *European Journal of Neuroscience*, vol. 15, no. 9, pp. 1499-1508, 2002.
- [Wendling 2005] Wendling, F., Hernandez, A., Bellanger, J. J., Chauvel, P., and Bartolomei, F., "Interictal to ictal transition in human temporal lobe epilepsy: insights from a

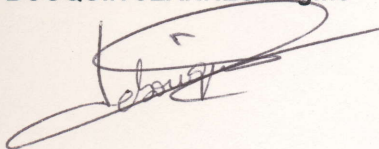
- computational model of intracerebral EEG," *Journal of Clinical Neurophysiology*, vol. 22, no. 5, pp. 343-356, 2005.
- [Wendling 2008] Wendling, F., and Chauvel, P., *Transition to ictal activity in temporal lobe epilepsy: insights from macroscopic models*, First ed., San Diego, CA, USA: Academic Press Inc, Elsevier, 2008.
- [Wendling 2010] Wendling, F., Chauvel, P., Biraben, A., and Bartolomei, F., "From Intracerebral EEG Signals to Brain Connectivity: Identification of Epileptogenic Networks in Partial Epilepsy," *Frontiers in Systems Neuroscience*, vol. 4, no. 154, pp. 1-13, 2010.
- [Wibral 2011a] Wibral, M., Rahm, B., Rieder, M., Lindner, M., Vicente, R., and Kaiser, J., "Transfer entropy in magnetoencephalographic data: Quantifying information flow in cortical and cerebellar networks," *Progress in Biophysics and Molecular Biology*, vol. 105, no. 1-2, pp. 80-97, 2011a.
- [Wibral 2011b] Wibral, M., Vicente, R., Priesemann, V., and Lindner, M., "TRENTOOL: an open source toolbox to estimate neural directed interactions with transfer entropy," *BMC Neuroscience*, vol. 12, no. Suppl 1, pp. P200, 2011b.
- [Wiener 1956] Wiener, N., *The Theory of Prediction*, New York: McGraw-Hill, 1956.
- [Wilson 1972] Wilson, H. R., and Cowan, J. D., "Excitatory and inhibitory interactions in localized populations of model neurons," *Biophysical Journal*, vol. 12, no. 1, pp. 1-24, 1972.
- [Wilson 1973] Wilson, H. R., and Cowan, J. D., "A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue," *Biological Cybernetics*, vol. 13, no. 2, pp. 55-80, 1973.
- [Winer 1986] Winer, S. L., "The role of exchange rate flexibility in the international transmission of inflation in long and shorter runs: Canada, 1953 to 1981," *Canadian Journal of Economics*, vol. 19, no. 1, pp. 62-86, 1986.
- [Winterhalder 2005] Winterhalder, M., Schelter, B., Hesse, W., Schwab, K., Leistriz, L., Klan, D., Bauer, R., Timmer, J., and Witte, H., "Comparison of linear signal processing techniques to infer directed interactions in multivariate neural systems," *Signal Processing*, vol. 85, no. 11, pp. 2137-2160, 2005.
- [Worrell 2004] Worrell, G. A., Parish, L., Cranstoun, S. D., Jonas, R., Baltuch, G., and Litt, B., "High-frequency oscillations and seizure generation in neocortical epilepsy," *Brain*, vol. 127, no. 7, pp. 1496-1506, 2004.
- [Yang 2010] Yang, C., Le Bouquin Jeannès, R., and Faucon, G., "Determining the flow direction of causal interdependence in multivariate time series," *18th European Signal Processing Conference (EUSIPCO)*, Aalborg, Denmark. pp. 636-640, Aug. 23-27 2010.
- [Yang 2011a] Yang, C., Le Bouquin Jeannès, R., Bellanger, J.-J., and Wendling, F., "Apport de l'entropie dans la détection du flux d'informations porté par des signaux modélisant des activités épileptiques," *XXIII Colloque Gretsi*, Bordeaux, France. pp. 1-4, Sep. 5-8 2011a.
- [Yang 2011b] Yang, C., Le Bouquin Jeannès, R., Faucon, G., and Shu, H., "Extracting information on flow direction in multivariate time series," *IEEE Signal Processing Letters*, vol. 18, no. 4, pp. 251-254, April 2011b.

- [Yang 2012] Yang, C., Le Bouquin Jeannès, R., Faucon, G., and Shu, H., "Detecting information flow direction in multivariate linear and nonlinear models," *Signal Processing*, 2012.
- [Zervakis 2011] Zervakis, M., Michalopoulos, K., Iordanidou, V., and Sakkalis, V., "Intertrial Coherence and Causal Interaction among Independent EEG Components," *Journal of Neuroscience Methods*, vol. 197, pp. 302-314, 2011.
- [Zetterberg 1978] Zetterberg, L. H., Kristiansson, L., and Mossberg, K., "Performance of a model for a local neuron population," *Biological Cybernetics*, vol. 31, no. 1, pp. 15-26, 1978.

VU :

Le Directeur de Thèse

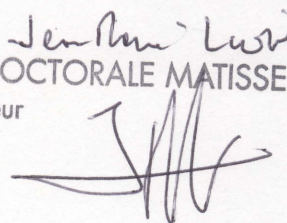
LE BOUQUIN JEANNES Régine



VU :

Le Responsable de l'École Doctorale

Le Responsable de l'École Doctorale
ÉCOLE DOCTORALE MATISSE
Le Directeur



DR UR/2012/S/7 n°86

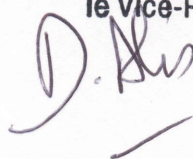
VU pour autorisation de soutenance

Rennes, le 20 juin 2012

Le Président de l'Université de Rennes 1


Guy CATHELINEAU

P/le Président et par délégation
le Vice-Président



VU après soutenance pour autorisation de publication :

Le Président de Jury,
(Nom et Prénom)

EDELINE Jean Marc


Contribution to effective connectivity analysis in epilepsy

Our work deals with effective connectivity to detect and quantify relations between cerebral structures involved in the initiation and the diffusion of epileptic seizures, aiming at establishing flow information propagation graphs. We study different approaches to answer two questions: (i) the identification of uni- and bi-directional relations, (ii) the discrimination between direct and indirect links. Firstly, we investigate the Granger causality index as well as its extended frequential and/or conditional versions, before exploiting a phase slope index and introducing a new indicator based on partial directed coherence. Then, we focus on transfer entropy selected as a nonlinear and nonparametric method computed from two signals. This method is considered in its conditional form to detect direct links taking into account the presence of a third signal. Since this technique is sensitive to calibration parameters such as the model order, a "greedy" strategy is proposed to optimize the order estimation based on the Bayesian information criterion. All approaches are evaluated and compared using Monte Carlo experiments on linear and nonlinear autoregressive models and also on physiology-based models and real signals recorded on an animal-model (guinea-pig) during a particular phase of a seizure corresponding to a narrowband tonic activity. Results on simulated signals allow us to establish coherent and consistent propagation graphs. For the real signals, without any ground-truth, which makes the assessment difficult, the use of surrogate data allows us to speculate a good behavior of our techniques and, for the three approaches tested, results appear coherent.

Contribution à l'analyse de la connectivité effective en épilepsie

Nos travaux s'inscrivent dans la problématique de l'identification de connectivité effective, et, donc, de graphes de propagation, pour détecter et quantifier les relations entre structures cérébrales impliquées dans l'initiation et la diffusion de crises épileptiques. Différentes approches sont envisagées pour apporter des éléments de réponse à deux questions essentielles, l'identification de liaisons unilatérales et/ou bilatérales et la mise en évidence de liens directs et/ou indirects. Nous nous intéressons d'abord à l'indice de causalité de Granger ainsi qu'à ses extensions fréquentielles et/ou conditionnées, avant d'explorer un indice de pente de phase et de proposer un nouvel indicateur construit sur la cohérence dirigée partielle. Puis, nous investissons une mesure d'entropie de transfert proposée comme méthode non linéaire et non paramétrique (à des horizons de prédiction près) dont nous considérons la forme conditionnelle dans une analyse multivariée. De par la subordination et la sensibilité de cette technique au choix de paramètres de calibration comme l'ordre des modèles, nous proposons une optimisation de cet ordre basée sur le critère d'information bayésien. Ces approches sont évaluées et comparées sur une base de signaux simulés par des processus vectoriels autorégressifs linéaires et non linéaires ainsi que sur des modèles physiologiques réalistes avant d'être appliquées sur des signaux réels enregistrés sur un modèle animal (cochon d'inde). En simulation, les résultats permettent d'établir des graphes de propagation cohérents et conformes aux modèles et, dans le cas réel, d'apprécier les variations de cette connectivité au cours du temps.